
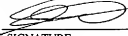


JC07 Rec'd PCT/PTO 19 NOV 2001

FORM PTO-1390 (REV. 9-2001)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER 066040-9675	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	
				09/980043	
INTERNATIONAL APPLICATION NO. PCT/US00/13637		INTERNATIONAL FILING DATE 18 May 2000 (18.05.2000)		PRIORITY DATE CLAIMED 21 May 1999 (21.05.1999)	
TITLE OF INVENTION METHOD FOR ENHANCING CELLULOSE AND MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS					
APPLICANT(S) FOR DO/EO/US Vincent L. Chiang, Luguang Wu, Chandrashekhar P. Joshi, and Daniel T. Carraway					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. 2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371. 3. <input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below. 4. <input checked="" type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (Article 31). 5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> has been communicated by the International Bureau. c. <input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)). a. <input type="checkbox"/> is attached hereto. b. <input type="checkbox"/> has been previously submitted under 35 U.S.C. 154(d)(4). 7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> have been communicated by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input checked="" type="checkbox"/> have not been made and will not be made. 8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)). 9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). 10. <input type="checkbox"/> An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).					
Items 11 to 20 below concern document(s) or information included:					
11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.					
12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.					
13. <input type="checkbox"/> A FIRST preliminary amendment.					
14. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.					
15. <input type="checkbox"/> A substitute specification.					
16. <input type="checkbox"/> A change of power of attorney and/or address letter.					
17. <input type="checkbox"/> A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.					
18. <input type="checkbox"/> A second copy of the published international application under 35 U.S.C. 154(d)(4).					
19. <input type="checkbox"/> A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).					
20. <input checked="" type="checkbox"/> Other items or information: Sequence Listing Statement and return receipt postcard.					
CERTIFICATION UNDER 37 CFR 1.10 I hereby certify that this correspondence is being deposited with the United States Postal Service on this date, November 19, 2001, as an envelope as "Express Mail Post Office to Addressee," mailing Label Number EL45386691US, addressed to the Box PCT, Commissioner for Patents, Washington, DC 20231 By: <i>Leslie Lindsay</i> Leslie Lindsay					

U.S. APPLICATION NO. (37 CFR 1.51) 097980043		INTERNATIONAL APPLICATION NO. PCT/US00/13637		ATTORNEY'S DOCKET NUMBER 066040-9675	
21. <input checked="" type="checkbox"/> The following fees are submitted: BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)): Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO. \$1040.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO \$890.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$740.00 International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$710.00 International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00 ENTER APPROPRIATE BASIC FEE AMOUNT =				CALCULATIONS PTO USE ONLY	
				\$	710.00
				\$	130.00
CLAIMS		NUMBER FILED		NUMBER EXTRA	
Total claims		46 - 20 =		26	
Independent claims		29 - 3 =		26	
				x	\$18.00
				x	\$84.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)				+	\$280.00
TOTAL OF ABOVE CALCULATIONS =				\$	3,492.00
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				+	\$ 1,746.00
SUBTOTAL =				\$	1,746.00
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				\$	0.00
TOTAL NATIONAL FEE =				\$	1,746.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +				\$	0.00
TOTAL FEES ENCLOSED =				\$	1,746.00
				Amount to be refunded:	\$
				charged:	\$
a. <input checked="" type="checkbox"/> A check in the amount of \$ <u>1,746.00</u> to cover the above fees is enclosed. b. <input type="checkbox"/> Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees. A duplicate copy of this sheet is enclosed. c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>13-3080</u> . A duplicate copy of this sheet is enclosed. d. <input type="checkbox"/> Fees are to be charged to a credit card. WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.					
NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO: Michael Best & Friedrich LLP 100 East Wisconsin Avenue, #3300 Milwaukee, WI 53202-4108					
		 23409 PATENT TRADEMARK OFFICE		 SIGNATURE Sara D. Vinarov NAME 48,524 REGISTRATION NUMBER	

JC13 Rec'd PCT/PTO 19 NOV 2001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
Group Art Unit – Unknown

In re

Patent Application of

Vincent L. Chiang, et al.

Serial No.: Unknown

Filed: November _____, 2001

Examiner: Unknown

"METHOD FOR ENHANCING CELLULOSE AND
MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS"

I, Leslie Lindsay, hereby certify that this correspondence is being deposited with the US Postal Service as "Express Mail Post Office to Addressee," mailing Label Number EL453986691US, Box PCT, Commissioner for Patents, Washington, DC 20231, on the date of my signature.


Signature
Date of Signature

SEQUENCE LISTING STATEMENT

Box PCT
Commissioner for Patents
Washington, DC 20231

Sir:

The Sequence Listing in this filing under 35 U.S.C. 371, is identical to the computer readable copy of the Sequence Listing in International Application No. PCT/US00/13637, filed May 18, 2000. In accordance with 37 CFR 1.821(e), please use the computer readable form filed in that application as the computer readable form for the instant application.

It is understood that the U.S. Patent and Trademark Office will make the necessary change in application number and filing date for the instant application. A paper copy of the Sequence Listing was included in the originally-filed specification in the instant application.

Respectfully submitted,



Sara D. Vinarov
Reg. No. 48,524

File No. 066040-9675

Michael Best & Friedrich LLP
One South Pinckney Street
P.O. Box 1806
Madison, WI 53701-1806
(608) 257-3501

Q:\CLIENT\066040\9675\B0106918.DOC

METHOD FOR ENHANCING CELLULOSE AND
MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS

5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Patent Application
Serial No. 60/135,280 filed 21 May, 1999.

10 STATEMENT REGARDING FEDERALLY SPONSORED
RESEARCH OR DEVELOPMENT

Not Applicable

FIELD OF THE INVENTION

15 This invention relates to polynucleotide molecules encoding cellulose
synthase, promoters of cellulose synthase and cellulose synthase polypeptides, methods for
genetically altering cellulose and lignin biosynthesis, and methods for improving strength
properties of juvenile wood and fiber in trees. The invention further relates to methods for
20 identifying regulatory elements in a cellulose synthase promoter and transcription factors
that bind to such regulatory elements, and to methods for augmenting expression of
polynucleotides operably linked to a cellulose synthase promoter.

BACKGROUND OF THE INVENTION

25 Lignin and cellulose are the two major building blocks of plant cell walls
that provide mechanical strength and rigidity. In plants, and especially in trees, these two
organic materials exist in a dynamic equilibrium conferring mechanical strength, water
transporting ability and protection from biotic and abiotic environmental stresses.
Normally, oven-dry wood contains 30 to 50% cellulose, 20 to 30% lignin and 20 to 30%
hemicellulose (Higuchi, 1997).

30 Proportions of lignin and cellulose are known to change with variation in
the natural environment. For example, during the development of compression wood in
conifers, the percentage of lignin increases from 30 to 40 %, and cellulose content
proportionally decreases from 40 to 30% (Timmell, 1986). Conversely, in angiosperm
tension wood the percentage of cellulose increases from 30 to 40%, while lignin content
35 decreases from 30 to 20% (Timmell, 1986).

It was recently discovered that the genetic down-regulation of a key tissue-
specific enzyme from the lignin biosynthesis pathway, 4CL, results in reduction of lignin

content by up to 45% in transgenic aspen trees (Hu *et al.*, 1999). This down-regulation is also associated with a 15% increase in the cellulose content. If the converse were true, *i.e.*, that increasing cellulose content by genetic up-regulation of cellulose biosynthesis results in reduction of lignin content, then the pulp yield could be increased. This would allow tremendous savings in chemical and energy costs during pulping because, for example, lignin must be degraded and removed during the pulping process.

Cellulose is a linear glucan consisting of β -D-1,4-linked glucose residues. It is formed by a cellulose synthase enzyme which catalyzes assembly of UDP-glucose units in plasma membrane complexes known as "particle rosettes" (Delmer and Amor, 1995). Cellulose synthase is thought to be anchored to the membrane by eight transmembrane binding domains to form the basis of the cellulose biosynthesis machinery in the plant cell wall (Pear *et al.*, 1996).

In higher plants, the glucan chains in cellulose microfibrils of primary and secondary cell walls are different in their degree of polymerization (Brown *et al.*, 1996). For example, secondary cell walls are known to contain cellulose having a high degree of polymerization, while in primary cell walls the degree of polymerization is lower. In another example, woody cell walls suffering from tension stress produce tension wood on the upper side of a bent angiosperm tree in response to the stress. In these cells, there are elevated quantities of cellulose which have very high crystallinity. The formation of highly crystalline cellulose is important to obtain a higher tensile strength of the wood fiber. Woody cell walls located at the under side of the same stem experience a compression stress, but do not produce highly crystalline cellulose. Such variation in the degree of polymerization in cell walls during development is believed to be due to different types of cellulose synthases for organizing glucose units into different paracrystalline arrays (Haigler and Blanton, 1996). Therefore, it would be advantageous to determine the molecular basis for the synthesis of highly crystalline cellulose so that higher yields of wood pulp having superior strength properties can be obtained from transgenic trees. Production of highly crystalline cellulose in transgenic trees would also markedly improve the mechanical strength properties of juvenile wood formed in normal trees. This would be a great benefit to the industry because juvenile wood is generally undesirable for solid wood applications because it has inferior mechanical properties.

Since the deposition of cellulose and lignin in trees is regulated in a compensatory fashion, genetic augmentation of cellulose biosynthesis might have a repressive effect on lignin deposition. Since the degree of polymerization and crystallinity may depend upon the type of cellulose synthase incorporated in the cellulose biosynthesis machinery, the expression of heterologous cellulose synthase or a UDP-glucose binding region thereof (*e.g.*, sweetgum protein expression in loblolly pine), could increase the quality of cellulose in transgenic plants. Over-expression of a heterologous cellulose

allow greater returns on investment by pulp and paper industries. Therefore, it would be advantageous to isolate and characterize genes from trees that are involved in cellulose biosynthesis in order to improve the properties of wood.

5

SUMMARY OF THE INVENTION

The present invention relates to polynucleotides comprising a nucleotide sequence that encodes a cellulose synthase, regulatory sequences, including a stress-inducible promoter, of the cellulose synthase, a cellulose synthase protein or a functional domain thereof and methods for augmenting cellulose biosynthesis in plants.

10

Thus, in one aspect, the invention provides a polynucleotide comprising a sequence that encodes a cellulose synthase, or a polynucleotide fragment thereof, the fragment encoding a functional domain of cellulose synthase, such as a UDP-glucose binding domain. The invention also provides a cellulose synthase or a functional domain or fragment thereof, including a UDP-glucose binding domain and at least one of eight transmembrane domains. The invention further provides a cellulose synthase promoter, or a functional fragment thereof, which fragment contains one or more mechanical stress response elements (MSRE).

15

20

In another aspect, the present invention is directed to a method of improving the quality of wood by altering the quantity of cellulose in plant cells, and optionally decreasing the content of lignin in the cell. The invention also relates to a method of altering the growth or the cellulose content of a plant by expressing an exogenous polynucleotide encoding a cellulose synthase or a UDP-glucose binding domain thereof in the plant. The invention further provides a method for causing a stress-induced gene expression in a plant cell by expressing a polynucleotide of choice using a stress-inducible cellulose synthase promoter.

25

In yet another aspect, the invention relates to a method for determining a mechanical stress responsive element (MSRE) in a cellulose synthase promoters and a method for identifying transcription factors that binds to the MSRE.

30

In a further aspect, the invention provides a method for altering (increasing or decreasing) i.e., regulating, the expression of a cellulose synthase in a plant by expressing an exogenous polynucleotide encoding a transcription factor having the property of binding a positive MSRE of a cellulose synthase promoter or by expressing an antisense polynucleotide encoding a transcription factor having the property of binding a negative MSRE to block the expression of the transcription factor.

35

Other aspects of the invention will be appreciated by a consideration of the detailed description of the invention drawings and appended claims.

DESCRIPTION OF THE DRAWINGS

Fig. 1 represents a nucleic acid sequence encoding a cellulose synthase from *Populus tremuloides* [SEQ ID NO: 1] and the protein sequence thereof [SEQ ID NO: 2].

Fig. 2 a-c (collectively referred to as Fig. 2) represent a Southern blot analysis of aspen genomic DNA probed with a fragment of the aspen cDNA represented in Fig. 1 under low (panel a) and high stringency conditions (panel b), and a northern blot analysis of the total aspen RNA from stem internodes using the same probe (panel c).

Fig. 3 a-d (collectively referred to as Fig. 3) represent *in situ* localization of the cellulose synthase gene transcripts as shown in the transverse sections from second (panel a), fourth (panel b), sixth (panel c) and fifth (panel d) internode.

Fig. 4 represents a nucleic acid sequence of the 5' region of aspen cellulose synthase gene including the promoter region and the 5' portion of the coding sequence [SEQ ID NO: 3].

Fig. 5 a-f (collectively referred to as Fig. 5) represents a histochemical analysis (panels a-d and f) and fluorescence microscopy (panel e) of transgenic tobacco for *GUS* gene expression driven by a cellulose synthase promoter of the invention.

Fig. 6 a-d (collectively referred to as Fig. 6) represents a histochemical analysis of *GUS* gene expression driven by aspen cellulose synthase promoter of the invention; tangential and longitudinal sections were harvested before bending (panel a), and 4 (panel b), 20 (panel c) and 40 (panel d) hours after bending and stained for *GUS* expression.

Fig. 7 represents a cDNA encoding cellulose synthase isolated from *Arabidopsis* [SEQ ID NO:4].

Fig. 8 represents an *Arabidopsis* cellulose synthase [SEQ ID NO:5] encoded by the cDNA represented in Fig. 7.

DETAILED DESCRIPTION OF THE INVENTION

All patents, patent applications and references cited in this specification are hereby incorporated herein by reference in their entirety. In case of any inconsistency, the present disclosure governs.

Definitions

The terms used in this specification generally have their ordinary meanings in the art, within the context of the invention, and in the specific context where each term is used. Certain terms are discussed below, or elsewhere in the specification, to provide additional guidance to the person of skill in the art in describing the compositions and methods of the invention and how to make and use them. It will be appreciated that the

same thing can be said in more than one way. Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein, nor is any special significance to be placed upon whether or not a term is elaborated or discussed herein. Synonyms for certain terms are provided. A recital of one or more synonyms does not exclude the use of other synonyms. The use of examples anywhere in this specification, including examples of any terms discussed herein, is illustrative only, and in no way limits the scope and meaning of the invention or of any exemplified term. Likewise, the invention is not limited to the preferred embodiments.

The term "plant" includes whole plants and portions of plants, including plant organs (e.g. roots, stems, leaves, etc.).

The term "angiosperm" refers to plants which produce seeds encased in an ovary. A specific example of an angiosperm is *Liquidambar styraciflua* (L.) [sweetgum].

The term "gymnosperm" refers to plants which produce naked seeds, that is, seeds which are not encased in an ovary. Specific examples of a gymnosperm include *Pinus taeda* (L.) [loblolly pine].

The term "polynucleotide" or "nucleic acid molecule" is intended to include double or single stranded genomic and cDNA, RNA, any synthetic and genetically manipulated polynucleotide, and both sense and anti-sense strands together or individually (although only sense or anti-sense strand may be represented herein). This includes single- and double-stranded molecules, i.e., DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases, for example thio-uracil, thio-guanine and fluoro-uracil.

An "isolated" nucleic acid molecule or polynucleotide refers to a component that is removed from its original environment (for example, its natural environment if it is naturally occurring). An isolated nucleic acid or polypeptide may contain less than about 50%, preferably less than about 75%, and most preferably less than about 90%, of the cellular components with which it was originally associated. A polynucleotide amplified using PCR so that it is sufficiently and easily distinguishable (on a gel, for example) from the rest of the cellular components is considered "isolated". The polynucleotides and polypeptides of the invention may be "substantially pure," i.e., having the highest degree of purity that can be achieved using purification techniques known in the art.

The term "hybridization" refers to a process in which a strand of nucleic acid joins with a complementary strand through base pairing. Polynucleotides are "hybridizable" to each other when at least one strand of one polynucleotide can anneal to a strand of another polynucleotide under defined stringency conditions. Hybridization requires that the two polynucleotides contain substantially complementary sequences;

depending on the stringency of hybridization, however, mismatches may be tolerated. Typically, hybridization of two sequences at high stringency (such as, for example, in an aqueous solution of 0.5X SSC at 65°C) requires that the sequences exhibit some high degree of complementarity over their entire sequence. Conditions of intermediate stringency (such as, for example, an aqueous solution of 2X SSC at 65°C) and low stringency (such as, for example, an aqueous solution of 2X SSC at 55°C), require correspondingly less overall complementarity between the hybridizing sequences. (1X SSC is 0.15 M NaCl, 0.015 M Na citrate.) As used herein, the above solutions and temperatures refer to the probe-washing stage of the hybridization procedure. The term "a polynucleotide that hybridizes under stringent (low, intermediate) conditions" is intended to encompass both single and double-stranded polynucleotides although only one strand will hybridize to the complementary strand of another polynucleotide.

A "sequence-conservative variant" is a polynucleotide that contains a change of one or more nucleotides in a given codon position, as compared with another polynucleotide, but the change does not result in any alteration in the amino acid encoded at that position.

A "function-conservative variant" is a polypeptide (or a polynucleotide encoding the polypeptide) having a given amino acid residue that has been changed without altering the overall conformation and function of the polypeptide, including, but not limited to, replacement of an amino acid with one having similar physico-chemical properties (such as, for example, acidic, basic, hydrophobic, and the like). Amino acids with have similar physico-chemical properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine, methionine or valine. Sequence- and function-conservative variants are discussed in greater detail below with respect to degeneracy of the genetic code.

A "functional domain" or a "functional fragment" refers to any region or portion of a protein or polypeptide or polynucleotide which is a region or portion of a larger protein or polynucleotide, the region or portion having the specific activity or specific function attributable to the larger protein or polynucleotide, e.g., a functional domain of cellulose synthase is the UDP-glucose binding domain.

The term "% identity" refers to the percentage of the nucleotides/amino acids of one polynucleotide/polypeptide that are identical to the nucleotides/amino acids of another sequence of polynucleotide/polypeptide as identified by program GAP from Genetics Computer Group Wisconsin (GCG) package (version 9.0) (Madison, WI). GAP uses the algorithm of Needleman and Wunsch (J. Mol. Biol. 48: 443-453, 1970) to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. When parameters required to run the above algorithm are

not specified, the default values offered by the program are contemplated. The following parameters are used by the GCG program GAP as default values (for polynucleotides): gap creation penalty:50; gap extension penalty:3; scoring matrix: nwsgapdna.cpm (local data file).

5 The "% similarity" or "% homology" between two polypeptide sequences is a function of the number of similar positions shared by two sequences on the basis of the scoring matrix used divided by the number of positions compared and then multiplied by 100. This comparison is made when two sequences are aligned (by introducing gaps if needed) to determine maximum homology. PowerBlast program, implemented by the
10 National Center for Biotechnology Information, can be used to compute optimal, gapped alignments. GAP program from Genetics Computer Group Wisconsin package (version 9.0) (Madison, WI) can also be used. GAP uses the algorithm of Needleman and Wunsch (J Mol Biol 48: 443-453, 1970) to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. When parameters
15 required to run the above algorithm are not specified, the default values offered by the program are contemplated. The following parameters are used by the GCG program GAP as default values (for polypeptides): gap creation penalty:12; gap extension penalty:4; scoring matrix:Blosum62.cpm (local data file).

The term "oligonucleotide" refers to a nucleic acid, generally of at least 10,
20 preferably at least 15, and more preferably at least 20 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule encoding a gene, mRNA, cDNA, or other nucleic acid of interest. Oligonucleotides can be labeled, e.g., with ³²P-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect
25 the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of *CelA*, or to detect the presence of nucleic acids encoding *CelA*. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a *CelA* DNA molecule. In still another embodiment, a library of oligonucleotides arranged on a solid
30 support, such as a silicon wafer or chip, can be used to detect various polymorphisms of interest. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

The term "coding sequence" refers to that portion of the gene that contains
35 the information for encoding a polypeptide. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic

sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences.

A "promoter" is a polynucleotide containing elements (e.g., a TATA box) which are capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Examples of promoters that can be used in the present invention include *PtCelAP*, 4CL-1 and 35S.

The term "constitutive promoter" refers to a promoter which typically does not require positive regulatory proteins to activate expression of an associated coding sequence, i.e., a constitutive promoter maintains some basal level of expression. A constitutive promoter is commonly used in creation of an expression cassette. An example of a constitutive promoter are 35S CaMV (Cauliflower Mosaic Virus), available from Clontech, Palo Alto, CA.

The term "inducible promoter" refers to the promoter which requires a positive regulation to activate expression of an associated coding sequence. An example of such a promoter is a stress-inducible cellulose synthase promoter from aspen described herein.

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence.

A "vector" is a recombinant nucleic acid construct, such as plasmid, phage genome, virus genome, cosmid, or artificial chromosome to which a polynucleotide of the invention may be attached. In a specific embodiment, the vector may bring about the replication of the attached segment, e.g., in the case of a cloning vector.

The term "expression cassette" refers to a polynucleotide which contains both a promoter and a protein coding sequence such that expression of a given protein is achieved upon insertion of the expression cassette into a cell.

A cell has been "transfected" by exogenous or heterologous polynucleotide when such polynucleotide has been introduced inside the cell. A cell has been "transformed" by exogenous or heterologous polynucleotide when the transfected polynucleotide effects a phenotypic change. Preferably, the transforming polynucleotide

should be integrated (covalently linked) into chromosomal DNA making up the genome of the cell.

"Exogenous" refers to biological material, such as a polynucleotide or protein, that has been isolated from a cell and is then introduced into the same or a different cell. For example, a polynucleotide encoding a cellulose synthase of the invention can be cloned from xylem cells of a particular species of tree, inserted into a plasmid and reintroduced into xylem cells of the same or different species. The species thus contains an exogenous cellulose synthase polynucleotide.

"Heterologous polynucleotide" refers to an exogenous polynucleotide not naturally occurring in the cell into which it is introduced.

"Homologous polynucleotide" refers to an exogenous polynucleotide that naturally exists in the cells into which it is introduced.

The present invention relates to isolation and characterization of polynucleotides encoding cellulose synthases from plants, especially trees, including full length or naturally occurring forms of cellulose synthases, functional domains, promoters and regulatory elements. Therefore, in accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook et al., 1989"); *DNA Cloning: A Practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* [B.D. Hames & S.J. Higgins eds. (1985)]; *Transcription And Translation* [B.D. Hames & S.J. Higgins, eds. (1984)]; *Animal Cell Culture* [R.I. Freshney, ed. (1986)]; *Immobilized Cells And Enzymes* [IRL Press, (1986)]; B. Perbal, *A Practical Guide To Molecular Cloning* (1984); F.M. Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

The present invention relates to a novel, full-length cellulose synthase gene (*CelA*), a novel stress inducible promoter of cellulose synthases (*CelAP*), and cellulose synthase proteins from trees, including UDP-glucose catalytic domains thereof. The invention enables the development of transgenic tree varieties having increased cellulose content, decreased lignin content and, therefore, improved wood fiber characteristics. Production of increased cellulose quantity and quality in multiple varieties of commercially relevant, transgenic forest tree species in operational production scenarios are further contemplated. The invention further provides a new experimental system for study of *CelA* gene expression and function in trees.

Polynucleotides encoding cellulose synthase and fragments thereof

The present invention relates to polynucleotides which comprise the nucleotide sequence that encodes cellulose synthase of the invention or a functional fragment thereof. In a preferred embodiment, the polynucleotide comprises the sequence encoding a tree cellulose synthase and most preferably, the sequence encoding a cellulose synthase from aspen. In one embodiment, a polynucleotide of the invention includes the entire cellulose synthase coding region, *e.g.*, nucleotides 69 to 3,005 of SEQ ID NO: 1. In another aspect of the invention, the polynucleotide encoding an *Arabidopsis* cellulose synthase is provided (see SEQ ID NO:4 and the translated protein of SEQ ID NO:5).

Also within the scope of the invention are fragments of the polynucleotides encoding cellulose synthase of the invention, which fragments encode at least one transmembrane domain and/or a UDP-glucose binding domain. For example, a polynucleotide comprising the nucleotides encoding a UDP-glucose binding domain of aspen cellulose synthase (*e.g.*, nucleotides 660 to 2250 of SEQ ID NO:1) or corresponding nucleotides of SEQ ID NO:4 are within the scope of the invention. The nucleotides encoding the UDP-glucose binding domain can be determined by, for example, alignment of protein sequences as described below.

The invention further relates to sequence conservative variants of the coding portion of SEQ ID NOS: 1 and 4.

Polynucleotides that hybridize under conditions of low, medium, and high stringency to SEQ ID NOS: 1 and 4, and their respective coding portions are also within the scope of the invention. Preferably, the polynucleotide that hybridizes to any of SEQ ID NOS: 1 and 4, or their respective coding portions, is about the same length as that sequence, for example, not more than about 10 to about 20 nucleotides longer or shorter.

In another embodiment of the invention, the hybridizable polynucleotide is at least 1500 nucleotides long, preferably at least 2500 nucleotides long and most preferably at least 3000 nucleotides long. In yet another embodiment, the hybridizable polynucleotide comprises the UDP-glucose binding domain as found in SEQ ID NO:1 or 4, or at least the conserved region QVLRW. Most preferably, the hybridizable polynucleotide has a UDP-glucose binding activity.

The polynucleotides that occur originally in nature may be isolated from the organisms that contain them using methods described herein or well known in the art. The non-naturally occurring polynucleotides may be prepared using various manipulations known in the field of recombinant DNA. For example, the cloned *Ce1A* polynucleotide can be modified according to methods described by Sambrook et al., 1989. The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the modified polynucleotides, for example, care should be taken to ensure that the modified

polynucleotide remains within the appropriate translational reading frame (if to be expressed) or uninterrupted by translational stop signals. As a further example, a *CelA*-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Preferably, such mutations enhance the functional activity of the mutated *CelA* polynucleotide. Any technique for mutagenesis known in the art can be used, including but not limited to, *in vitro* site-directed mutagenesis (Hutchinson, C., et al., 1978, J. Biol. Chem. 253:6551; Zoller and Smith, 1984, DNA 3:479-488; Oliphant et al., 1986, Gene 44:177; Hutchinson et al., 1986, Proc. Natl. Acad. Sci. U.S.A. 83:710), use of TAB linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis (see Higuchi, 1989, "Using PCR to Engineer DNA", in *PCR Technology: Principles and Applications for DNA Amplification*, H. Erlich, ed., Stockton Press, Chapter 6, pp. 61-70).

The polynucleotides of the present invention may be introduced into various vectors adapted for plant or non-plant replication. These are well known in the art, thus, choice, construction and use of such vectors is well within the skill of a person skilled in the art. Possible vectors include, but are not limited to, plasmids or modified viruses of plants, but the vector system must be compatible with the host cell used. An example of a suitable vector is Ti plasmid. The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. An expression cassette containing cellulose synthase or recombinant molecules thereof can be introduced into host cells via silicon carbide whiskers, transformed protoplasts, transformation, e.g., *Agrobacterium* vectors (discussed below), electroporation, infection, etc., so that many copies of the gene sequence are generated. Preferably, the cloned gene is contained on a shuttle vector plasmid, which provides for expansion in a cloning cell, e.g., *E. coli*, and facile purification for subsequent insertion into an appropriate expression cell line, if such is desired. For example, a shuttle vector, which is a vector that can replicate in more than one type of organism, can be prepared for replication in both *E. coli* and *Saccharomyces cerevisiae* by linking sequences from an *E. coli* plasmid with sequences from the yeast 2m plasmid.

Transgenic plants containing the polynucleotides described herein are also within the scope of the invention. Methods for introducing exogenous polynucleotides

into plant cells and regenerating transgenic plants are well known. Some are provided below.

In one embodiment, to introduce a plasmid containing a *CelA* coding sequence or promoter of the invention into a plant, a 1:1 mixture of plasmid DNA containing a selectable marker expression cassette and plasmid DNA containing a cellulose synthase expression cassette is precipitated with gold to form microprojectiles. The microprojectiles are rinsed in absolute ethanol and aliquots are dried onto a suitable macrocarrier such as the macrocarrier available from BioRad in Hercules, CA. Prior to bombardment, embryonic tissue is preferably desiccated under a sterile laminar-flow hood. The desiccated tissue is transferred to semi-solid proliferation medium. The prepared microprojectiles are accelerated from the macrocarrier into the desiccated target cells using a suitable apparatus such as a BioRad PDS-1000/HE particle gun. In a preferred method, each plate is bombarded once, rotated 180 degrees, and bombarded a second time. Preferred bombardment parameters are 1350 psi rupture disc pressure, 6 mm distance from the rupture disc to macrocarrier (gap distance), 1 cm macrocarrier travel distance, and 10 cm distance from macrocarrier stopping screen to culture plate (microcarrier travel distance). Tissue is then transferred to semi-solid proliferation medium containing a selection agent, such as hygromycin B, for two days after bombardment.

Cellulose synthase protein and fragment thereof

A cellulose synthase of the invention is a plant protein that contains a catalytic subunit which has UDP-glucose binding activity for the synthesis of glucan from glucose, and eight transmembrane domains for localizing the cellulose synthase to the cell membrane. The cellulose synthase of the invention has eight transmembrane binding domains; two at the amino terminal and six at the carboxyl terminal. The UDP-glucose binding domain is located between transmembrane domains two and three. Examples of this protein structure are seen in the aspen cellulose synthase as well as in those of RSW1 and GhCelA. The location of the transmembrane domain may be identified as described below and as exemplified in the Example. Preferably, the cellulose synthase of the invention has an amino acid sequence of a tree cellulose synthase.

In one embodiment, the cellulose synthase protein of the invention is isolated from aspen. Aspen cellulose synthase contains about 978 amino acids and has a molecular weight of about 110 KDa and a pI of about 6.58. In one embodiment, the aspen cellulose synthase has the amino acid sequence of SEQ ID NO:2 as represented in Fig. 1. In another aspect, the invention relates to cellulose synthase of SEQ ID NO: 5.

The invention further relates to fragments of plant cellulose synthases, such as fragments containing at least one transmembrane region and/or a UDP-glucose binding

domain. The transmembrane regions may be identified as described in the Example by using the method of Hoffman and Stoffel (1993).

The cellulose synthase fragment containing the UDP-glucose binding domain is functional without the presence of the rest of the protein. This separable activity is as shown in the Example. This result was surprising and unexpected because previously identified UDP-glucose binding domains were not known to be functional when isolated from other portions of the protein. Thus, a fragment of any cellulose synthase (such as *PtCelA*, *RSW1*, *GhCelA* and SEQ ID NO:5) that contains a UDP-glucose binding domain and is independently functional is within the scope of the invention. The function of the UDP-glucose binding domain may be determined using the assay described in the Example. The UDP-glucose binding domain of the invention is located between the second and third transmembrane region of the cellulose synthase and has conserved amino acid sequences for UDP-glucose binding, such as the sequence QVLRW and conserved D residues. The UDP-glucose binding domain and the conserved regions therein may be located in a cellulose synthase using the guidance of the present specification and the general knowledge in the art, for example Brown, 1996. In one embodiment, the UDP-glucose binding domain and the conserved regions therein may be identified by comparing the amino acid sequence of cellulose synthase of interest with the amino acid sequence of aspen cellulose synthase using the algorithms described in the specification or generally known in the art. For example, the UDP-glucose binding domain of SEQ ID NO:2 is in the position amino acids 220 to 749. The conserved QVLRW sequence is located at positions 715-719 of SEQ ID NO:2.

Polypeptides having at least 75%, preferably at least 85% and most preferably at least 95% similarity to the amino acid sequence of SEQ ID NO: 2, amino acids 220-749 of SEQ ID NO:2, SEQ ID NO:5 or its UDP-glucose binding domain using Power Blast or GAP algorithm described above. In a preferred embodiment, these polypeptides are of about the same length as the polypeptide of SEQ ID NO: 2 or amino acids 220-749 of SEQ ID NO:2. For example, the polypeptide may be from about 2-3 to about 5-7 and to about 10-15 amino acids longer or shorter. In another embodiment, the polypeptides described in this paragraph are not originally found (*i.e.*, naturally occurring) in *Arabidopsis* or cotton. These polypeptides may be prepared by, for example, altering the nucleic acid sequence of a cloned polynucleotide encoding the protein of SEQ ID NO:2 or SEQ ID NO:5 using the methods well known in the art.

Function conservative variants of cellulose synthase are also within the scope of the invention and can be prepared by altering the sequence of a cloned polynucleotide encoding cellulose synthase or fragments thereof. Conventional methods used in the art can be used to make substitutions, additions or deletions in one or more amino acids, to provide functionally equivalent molecules. For example, a function

conservative variant that has substitutions, deletions and/or additions in the amino and/or carboxyl terminus of the protein, outside of the UDP-glucose binding domain is within the scope of the invention. Preferably, variants are made that have enhanced or increased functional activity relative to native cellulose synthase. Methods of directed evolution can be used for this purpose.

The invention also includes function conservative variants which include altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a conservative amino acid substitution. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity, which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. Amino acids containing aromatic ring structures are phenylalanine, tryptophan, and tyrosine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Such alterations will not be expected to affect apparent molecular weight as determined by polyacrylamide gel electrophoresis, or isoelectric point. Particularly preferred substitutions are: (i) Lys for Arg and vice versa such that a positive charge may be maintained; (ii) Glu for Asp and vice versa such that a negative charge may be maintained; (iii) Ser for Thr such that a free -OH can be maintained; and (iv) Gln for Asn such that a free CONH₂ can be maintained. Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β -turns in the protein's structure.

The cellulose synthase of the invention can be isolated by expressing a cloned polynucleotide encoding the cellulose synthase as well as using direct protein purification techniques. These methods will be apparent to those of skill in the art.

Polynucleotides containing cellulose synthase promoter

The present invention further relates to a cellulose synthase promoter. The promoter is a stress-inducible promoter and may be used to synthesize greater quantities of high crystalline cellulose in plant, and preferably in trees. This permits an increase in the

proportion of cellulose in transgenic plants, greater strength of juvenile wood and fiber, and acceleration of overall growth rate.

In one embodiment, the promoter of the invention is from aspen and is represented in Figure 4. The promoter sequence is located within the region of nucleotides 1-840 of SEQ ID NO:3. A person of skill in the art will appreciate that not the entire sequence is required for the promoter function and can easily identify the critical regions by looking for conserved boxes and doing routine deletion analysis. Thus, functional fragments of SEQ ID NO:1 are within the scope of the invention.

Polynucleotides that hybridize under conditions of low, medium, and high stringency to SEQ ID NO:3, and its non-coding portion are also within the scope of the invention. The hybridizable polynucleotide may be about the same length as the sequence to which it hybridizes, for example, not more than about 10 to about 20 nucleotides longer or shorter. In another embodiment, the hybridizable polynucleotide is at least about 200 nucleotides long, preferably at least about 400 nucleotides long and most preferably at least 500 nucleotides long. In yet another embodiment, the hybridizable polynucleotide comprises at least one MSRE element identified according to the method described below.

A cellulose synthase promoter of the invention typically provides tissue-specific gene regulation in xylem, but also permits up-regulation of gene expression in other tissues as well, *e.g.*, phloem under tension stress. Furthermore, expression of cellulose synthase is localized to an area of the plant under stress.

This stress-inducible phenomenon is regulated by positive and negative mechanical stress response elements (MSREs). These MSREs upregulate (positive) or downregulate (negative) the expression of a cellulose synthase polynucleotide under stress conditions through binding of transcription factors. MSRE-regulated expression of cellulose synthase permits synthesis of cellulose with high crystallinity.

The MSREs of cellulose synthase can be modified or employed otherwise in methods to regulate expression of a polynucleotide, including a cellulose synthase, operatively linked to a promoter containing an MSRE in response to mechanical stress (*e.g.*, tension or compression) to a transgenic plant.

Negative MSREs of a cellulose synthase promoter can be modified, removed or blocked to improve expression of a cellulose synthase, and thereby increase cellulose production and improve wood quality. Alternatively, positive MSREs can be removed or blocked to decrease expression of a cellulose synthase, which decreases cellulose production and increases lignin deposition. This is useful for increasing the fuel value of wood because lignin has a higher BTU value than cellulose. Moreover, a modified cellulose synthase promoter can be operatively linked to a polynucleotide of interest to control its expression upon mechanical stress to a plant harboring it.

The location of MSRE elements in the SEQ ID NO:3 may be identified, for example, using promoter deletion analysis, DNase Foot Print Analysis, and Southwestern screening of an expression library for an MSRE. In one embodiment, cellulose synthase promoter that has one or more portions deleted, and is operatively linked to a reporter sequence, is introduced into a plant or a plant cell. A positive MSRE is detected by observing no relative change or increase in the amount of reporter in a transgenic plant or tissue, *e.g.*, phloem after inducing a stress to the plant, and a negative MSRE is detected by observing increases in the amount of reporter in the plant in the absence of any stress to the plant. A positive element is detected when by removing it, GUS expression goes down and by adding it kept at the same level or more. The negative element does not support, or suppresses, expression of GUS and by removing it, normal or enhanced GUS expression is observed as compared to when negative element is present.

Manipulation of a MSRE binding sites and/or providing transcription factors that bind thereto, provides a mechanism to continuously produce high crystalline cellulose in woody plant cell walls of transgenic plants. For example, one having ordinary skill in the art can delete or block negative MSRE elements, or provide cDNA encoding protein(s) that bind the positive MSREs, to enable constitutive expression of a cellulose synthase without the requirement of a mechanical stress. The increased cellulose synthase, and therefore, increased cellulose content, can improve the strength properties of juvenile wood and fiber. It is also contemplated that the positive MSREs can be deleted or blocked, or cDNA in an antisense direction, which in the sense direction encodes a protein that binds a positive MSRE, can be provided, to reduce cellulose synthase activity and decrease cellulose production.

Method of Isolating Polynucleotides Encoding Cellulose Synthase

The invention further relates to identifying and isolating polynucleotides encoding cellulose synthase in plants, *e.g.*, trees, (in addition to those polynucleotides provided in the Example and represented in Fig. 1 and Fig. 7). These polynucleotides may be used to manipulate expression of cellulose synthase with an objective to improve the cellulose content and properties of wood.

The method comprises identifying a nucleic acid fragment containing a sequence encoding cellulose synthase or a portion thereof by using a fragment of SEQ ID NOS:1 or 4 as a probe or a primer. Once identified, the nucleic acid fragment containing a sequence encoding cellulose synthase or a portion thereof is isolated.

Polynucleotides encoding cellulose synthases of the invention, whether genomic DNA, cDNA, or fragments thereof, can be isolated from many sources, particularly from cDNA or genomic libraries from plants, preferably trees (*e.g.* aspen, sweetgum, loblolly pine, eucalyptus, and other angiosperms and gymnosperms).

Molecular biology methods for obtaining polynucleotides encoding a cellulose synthase are well known in the art, as described above (*see, e.g., Sambrook et al., 1989, supra*).

Accordingly, cells from any species of plant can potentially serve as a nucleic acid source for the molecular cloning of a polynucleotide encoding a cellulose synthase of the invention. The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g., a DNA "library"*), and preferably is obtained from a cDNA library prepared from tissues with high level expression of a cellulose synthase (*e.g., xylem tissue, since cells in this tissue evidence very high levels of expression of CelA*), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from a desired cell (*see, for example, Sambrook et al., 1989, supra; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II*). Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will not contain intron sequences. Whatever the source, a polynucleotide should be molecularly cloned into a suitable vector for its propagation.

In another embodiment for the molecular cloning of a polynucleotide encoding a cellulose synthase of the invention from genomic DNA, DNA fragments are generated from a genome of interest, such as from a plant, or more particularly a tree genome, part of which will correspond to a desired polynucleotide. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNase in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques, including but not limited to, agarose and polyacrylamide gel electrophoresis and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing a desired *CelA* sequence may be accomplished in a number of ways. For example, if an amount of a portion of a *CelA* sequence or its specific RNA, or a fragment thereof, is available and can be purified and labeled, the generated DNA fragments may be screened by nucleic acid hybridization to a labeled probe (Benton and Davis, 1977, *Science* 196:180; Grunstein and Hogness, 1975, *Proc. Natl. Acad. Sci. U.S.A.* 72:3961). For example, a set of oligonucleotides corresponding to the partial amino acid sequence information obtained for a *CelA* protein from trees can be prepared and used as probes for DNA encoding cellulose synthase, or as primers for cDNA or mRNA (*e.g., in combination with a poly-T primer for RT-PCR*). Preferably, a fragment is selected that is highly unique to a cellulose synthase of the invention, such as the UDP-glucose binding regions. Those DNA fragments with substantial homology to the probe will hybridize. As noted above, the greater the degree of homology, the more stringent

hybridization conditions can be used. In a specific embodiment, stringency hybridization conditions can be used to identify homologous *CelA* sequences from trees or other plants.

Thus, in one embodiment, a labeled cellulose synthase cDNA from, e.g., *Populus tremuloides* (*PtCelA*), can be used to probe a library of genes or DNA fragments from various species of plants, especially angiosperm and gymnosperm, to determine whether any bind to a *CelA* of the invention. Once genes or fragments are identified, they can be amplified using standard PCR techniques, cloned into a vector, e.g., pBluescript vector (StrataGene of LaJolla, CA), and transformed into a bacteria, e.g., DH5 α *E. coli* strain (Gibco BRL of Gaithersburg, MD). Bacterial colonies are typically tested to determine whether any contains a cellulose synthase-encoding nucleic acid. Once a positive clone is identified through binding, it is sequenced from an end, preferably the 3' end.

cDNA libraries can be constructed in various hosts, such as lambda ZAPII, available from Stratagene, LaJolla, CA, using poly(A) +rRNA isolated from aspen xylem, according to the methods described by Bugos *et al.* (Biotechniques 19:734-737, 1995). The above mentioned probes are used to assay the aspen cDNA library to locate cDNA which codes for enzymes involved in production of cellulose synthases. Once a cellulose synthase sequence is located, it is then cloned and sequenced according to known methods in the art.

Further selection can be carried out on the basis of the properties of the gene, *e.g.*, if the gene encodes a protein product having the isoelectric, electrophoretic, hydropathy plot, amino acid composition, or partial amino acid sequence of a cellulose synthase protein of the invention, as described herein. Thus, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones or DNA clones which hybrid-select the proper mRNAs can be used to produce a protein that has similar properties known for cellulose synthases of the invention. Such properties may include, for example, similar or identical electrophoretic migration patterns, isoelectric focusing or non-equilibrium pH gel electrophoresis behavior, proteolytic digestion maps, hydropathy plots, or functional properties (such as isolated, functional UDP-glucose binding domains).

A cellulose synthase polynucleotide of the invention can also be identified by mRNA selection, *i.e.*, by nucleic acid hybridization followed by *in vitro* translation. In this procedure, nucleotide fragments are used to isolate complementary mRNAs by hybridization. Such DNA fragments may represent available, purified *CeLa* DNA, or may be synthetic oligonucleotides designed from the partial amino acid sequence information. Functional assays (*e.g.*, UDP-glucose activity) of the *in vitro* translation products of the products of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments, that contain the desired sequences.

A radiolabeled *CelA* cDNA can be synthesized using a selected mRNA as a template. The radiolabeled mRNA or cDNA may then be used as a probe to identify homologous *CelA* DNA fragments from amongst other genomic DNA fragments.

It will be appreciated that other polynucleotides, in addition to a *CelA* of the invention can be operatively linked to a *CelA* promoter to control expression of the polynucleotide upon application of a mechanical stress.

Expression of *CelA* Polypeptides

The nucleotide sequence coding for *CelA*, or a functional fragment, derivative or analog thereof, including chimeric proteins, can be inserted into an appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Preferably, an expression vector includes an origin of replication. The elements are collectively termed herein a "promoter." Thus, a nucleic acid encoding *CelA* of the invention can be operatively associated with a promoter in an expression vector of the invention. Both cDNA and genomic sequences can be cloned and expressed under control of such regulatory sequences. The necessary transcriptional and translational signals can be provided on a recombinant expression vector, or they may be supplied by the native gene encoding *CelA* and/or its flanking regions.

In addition to a *CelAP*, expression of cellulose synthase can be controlled by any promoter/enhancer element known in the art, but these regulatory elements must be functional in the host selected for expression. Promoters which may be used to control *CelA* polynucleotide expression include, constitutive, development-specific and tissue-specific. Examples of these promoters include 35S Cauliflower Mosaic Virus, terminal flower and 4CL-1. Thus, there are various ways to alter the growth of a plant using different promoters, depending on the needs of the practitioner.

The nucleotide sequence may be inserted in a sense or antisense direction depending on the needs of the practitioner. For example, if augmentation of cellulose biosynthesis is desired then polynucleotides encoding, *e.g.*, cellulose synthase, can be inserted into the expression vector in the sense direction to increase cellulose synthase production and thus cellulose biosynthesis. Alternatively, if it is desired that cellulose biosynthesis is reduced or lignin content is increased, then polynucleotides encoding, *e.g.*, cellulose synthase, can be inserted in the antisense direction so that upon transcription the antisense mRNA hybridizes to other complementary transcripts in the sense orientation to prevent translation. In other embodiments, the polynucleotide encodes a UDP-glucose binding domain and is used in a similar manner as described.

A recombinant *CelA* protein of the invention, or functional fragment, derivative, chimeric construct, or analog thereof, may be expressed chromosomally, after

integration of the coding sequence by recombination. In this regard, any of a number of amplification systems for plants may be used to achieve high levels of stable gene expression, as discussed above. Any of the methods previously described for the insertion of DNA fragments into a cloning vector may be used to construct expression vectors
5 containing a gene consisting of appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include *in vitro* recombinant DNA and synthetic techniques and *in vivo* recombination (genetic recombination).

Expression vectors containing a nucleic acid encoding a *CelA* of the invention can be identified by four general approaches: (a) PCR amplification of the
10 desired plasmid DNA or specific mRNA, (b) nucleic acid hybridization, (c) presence or absence of selection marker gene functions, (d) analyses with appropriate restriction endonucleases, and (e) expression of inserted sequences. In the first approach, the nucleic acids can be amplified by PCR to provide for detection of the amplified product. In the second approach, the presence of a foreign gene inserted in an expression vector can be
15 detected by nucleic acid hybridization using probes comprising sequences that are homologous to an inserted marker gene. In the third approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "selection marker" gene functions (*e.g.*, β -glucuronidase activity, resistance to antibiotics, transformation phenotype, etc.) caused by the insertion of foreign genes in the
20 vector. In another example, if the nucleic acid encoding *CelA* is inserted within the "selection marker" gene sequence of the vector, recombinants containing the *CelA* insert can be identified by the absence of the *CelA* gene function. In the fourth approach, recombinant expression vectors are identified by digestion with appropriate restriction enzymes. In the fifth approach, recombinant expression vectors can be identified by
25 assaying for the activity, biochemical, or immunological characteristics of the gene product expressed by the recombinant, provided that the expressed protein assumes a functionally active conformation.

After a particular recombinant DNA molecule is identified and isolated, several methods known in the art may be used to propagate it. Once a suitable host system
30 and growth conditions are established, recombinant expression vectors can be propagated and prepared in quantity. As previously explained, the expression vectors which can be used include, but are not limited to those vectors or their derivatives described above.

Vectors are introduced into the desired host cells by methods known in the art, *e.g.*, *Agrobacterium*-mediated transformation (described in greater detail below),
35 transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, lipofection (lysosome fusion), use of a gene gun, or a DNA vector transporter (see, *e.g.*, Wu et al., 1992, J. Biol. Chem. 267:963-967; Wu and

Any non-tumorigenic *A. tumefaciens* strain harboring a disarmed Ti 35 plasmid may be used in the method of the invention. Any *Agrobacterium* system may be used. For example, Ti plasmid/binary vector system or a cointegrative vector system with one Ti plasmid may be used. Also, any marker gene or polynucleotide conferring the ability to select transformed cells, callus, embryos or plants and any other gene, such as,

for example, a gene conferring resistance to a disease, or one improving cellulose content, may also be used. Any promoter desired can be used, such as, for example, a *PtCelAP* of the invention, and those promoters as described above. A person of ordinary skill in the art can determine which markers and genes are used depending on particular needs.

5 For purposes of the present invention, "transformed" or "transgenic" means that at least one marker gene or polynucleotide conferring selectable marker properties is introduced into the DNA of a plant cell, callus, embryo or plant. Additionally, any gene may also be introduced.

To increase the infectivity of the bacteria, *Agrobacterium* is cultured in
10 low-pH induction medium, i.e., any bacterium culture media with a pH value adjusted to from 4.5 to 6.0, most preferably about 5.2, and at low temperature such as for example about 19-30°C, preferably about 21-26°C. The conditions of low-pH and low temperature are among the well-defined critical factors for inducing virulence activity in *Agrobacterium* (e.g., Altmorbe *et al.*, Mol. Plant-Microbe. Interac. 2: 301, 1989; Fullner *et al.*, Science 273: 1107, 1996; Fullner and Nester, J. Bacteriol. 178: 1498, 1996).

The bacteria is preconditioned by coculturing with wounded tobacco leaf extract (prepared according to methods known generally known in the art) to induce a high level of expression of the *Agrobacterium vir* genes. Prior to inoculation of plant somatic embryos, *Agrobacterium* cells can be treated with a tobacco extract prepared from
20 wounded leaf tissues of tobacco plants grown *in vitro*. To achieve optimal stimulation of the expression of *Agrobacterium vir* genes by wound-induced metabolites and other cellular factors, tobacco leaves can be wounded and pre-cultured overnight. Culturing of bacteria in low pH medium and at low temperature can be used to further enhance the bacteria *vir* gene expression and infectivity. Preconditioning with tobacco extract and the
25 *vir* genes involved in the T-DNA transfer process are generally known in the art.

Agrobacterium treated as described above is then cocultured with a plant tissue explant, such as for example zygotic and/or somatic embryo tissue. Non-zygotic (i.e., somatic) or zygotic tissues can be used. Any plant tissue may be used as a source of explants. For example, cotyledons from seeds, young leaf tissue, root tissues, parts of
30 stems including nodal explants, and tissues from primary somatic embryos such as the root axis may be used. Generally, young tissues are a preferred source of explants.

The invention also relates to methods of altering the growth of a plant by expressing the polynucleotide of the invention, which as a result alters the growth of the plant. The polynucleotide used in the method may be a homologous polynucleotide or a
35 heterologous polynucleotide and are described in detail above. For example, both full-length and UDP-glucose binding region containing fragments may be expressed. Additionally, depending on the aim of the method, the polynucleotide may be introduced into the plant in the sense or in the antisense orientation. Any suitable promoter may be

used to provide expression. The promoter or a functional fragment thereof is operatively linked to the polynucleotide. The promoter may be a constitutive promoter, a tissue-specific promoter or a development-specific plant promoter. Examples of suitable promoters are Cauliflower Mosaic Virus 35S, 4CL, cellulose synthase promoter, *PrCelAP* and terminal flower promoter.

The invention further relates to a method of altering the cellulose content in a plant by expressing the polynucleotide of the invention as described above. The method may be used to increased the ratio of cellulose to lignin in the plant that have an exogenous polynucleotide of the invention introduced therein.

The invention further relates to a method for altering expression of a cellulose synthase in a plant cell by introducing into the cell a vector comprising a polynucleotide of the invention and expressing the polynucleotide. The polynucleotides and promoters described above may be used.

A method for causing stress-induced gene expression in a plant cell is also within the scope of the invention. The method comprises (i) introducing into the plant or a plant cell an expression cassette comprising a cellulose synthase promoter or a functional fragment thereof or providing a plant or a plant cell that comprises the expression cassette (The promoter of the cassette is operatively linked to a coding sequence of choice.); and (ii) applying mechanical stress to the plant to induce expression of the desired coding sequence.

A method for determining a positive mechanical stress responsive element (MSRE) in a cellulose synthase promoter is also within the scope of the invention and comprises (i) making serial deletions in the cellulose synthase promoter, such as for example, SEQ ID NO:3; (ii) introducing the deletion linked to a polynucleotide encoding a reporter sequence into a plant cell, and (iii) detecting a decrease in the amount of reporter in the plant after inducing a stress to the plant. Similarly, a method for determining a negative MSRE in a cellulose synthase promoter is provided. It comprises (i) making serial deletions in the cellulose synthase promoter, such as for example, SEQ ID NO:3; (ii) introducing the deletion linked to a polynucleotide encoding a reporter sequence into a plant cell, and (iii) detecting an increase in the amount of reporter in the plant after inducing a stress to the plant.

The following methods are also within the scope of the invention: a method for expressing cellulose synthase in a tissue-specific manner comprising transforming a plant with a tissue specific promoter operatively linked to a polynucleotide encoding a cellulose synthase; a method for inducing expression of a cellulose synthase in a plant comprising introducing into a plant a cDNA encoding a protein that binds to a positive MSRE of a cellulose synthase promoter, thereby resulting in increased expression of cellulose in the plant, wherein the binding to the positive MSRE results in expression of

a cellulose synthase; a method for reducing expression of a cellulose synthase comprising introducing into a plant a cDNA in an antisense orientation, wherein the cDNA in a sense orientation encodes a protein that binds to a positive MSRE and results in expression of a cellulose synthase; a method for increasing cellulose biosynthesis in a plant comprising introducing into a plant a cDNA encoding a protein that binds to a positive MSRE of a cellulose synthase promoter, whereby binding of the protein to the positive MSRE results in expression of a cellulose synthase, and A method for reducing cellulose biosynthesis in a plant comprising introducing into a plant a cDNA in an antisense orientation, wherein the cDNA in a sense orientation encodes a protein that binds to a positive MSRE of a cellulose synthase promoter.

EXAMPLE

Molecular cloning of cellulose synthase

This Example describes the first tree cellulose synthase cDNA (*PtCela*, GenBank No. AF072131) cloned from developing secondary xylem of aspen trees using RSW1 cDNA.

Prior to the present invention, only partial clones of cellulose synthases from crop species and cotton *GhCela* have been discovered, which have significant homology to each other. The present inventors have discovered and cloned a new full-length cellulose synthase cDNA, *AraxCela* (GenBank No. AF062485) (Fig. 7, [SEQ ID NO: 4]), from an *Arabidopsis* primary library. *AraxCela* is a new member of cellulose synthase and shows 63-85% identity and 72-90% similarity in amino acid sequence with other *Arabidopsis Cela* members.

Another cellulose synthase was cloned in aspen using a ³²P-labeled 1651-bp long *EcoRI* fragment of *Arabidopsis Cela* cDNA, which encodes a centrally located UDP-glucose binding domain, was used as a probe to screen about 500,000 pfu of a developing xylem cDNA library from aspen (*Populus tremuloides*) (Ge and Chiang, 1996). Four positive clones were obtained after three rounds of plaque purification. Sequencing the 3' ends of these four cDNAs showed that they were identical clones. The longest cDNA clone was fully sequenced and determined to be a full-length cDNA having a 3232 bp nucleotide sequence (Fig. 1) [SEQ ID NO: 1], which encodes a protein of 978 amino acids [SEQ ID NO: 2].

Characterization of a cellulose synthase from aspen

The first AUG codon of *PtCela* was in the optimum context for initiation of transcription on the basis of optimal context sequence described by Joshi (1987a) and Joshi *et al.* (1997). A putative polyadenylation signal (AATACA) was found 16 bp upstream of a polyadenylated tail of 28 bp, which is similar to the proposed plant structure (Joshi, 1987b). The 5' untranslated leader was determined to have 68 bp and the 3'

untranslated trailer was 227 bp. Both of these regions have a typical length observed in many plant genes (Joshi, 1987a and Joshi, 1987b). This cDNA clone exhibited 90% amino acid sequence similarity with cellulose synthase from cotton (*GhCelA*), and 71% with cellulose synthase from *Arabidopsis* (*RSW1*), suggesting that this particular tree homolog also encodes a cellulose synthase.

The full length cDNA was designated *PrCelA*, and encodes a 110,278 Da polypeptide having an isoelectric point (pI) of 6.58 and 8 charged molecules. The hydropathy curve indicated that this particular cellulose synthase has eight transmembrane binding domains; two at the amino terminal and six at the carboxyl terminal, using the method of Hoffman and Stoffel (1993). This protein structure is analogous to those of *RSW1* and *GhCelA*. All of the conserved domains for UDP-glucose binding, such as QVLRW and conserved D residues, are also present in a cellulose synthase of the invention, e.g., *PrCelA* (Brown *et al.*, 1996). Thus, based on sequence and molecular analyses, it was concluded that *PrCelA* encodes a catalytic subunit which, like *RSW1* in *Arabidopsis*, is essential for the cellulose biosynthesis machinery in aspen.

In situ localization of *PrCelA* mRNA transcripts along the developmental gradient defined by stem primary and secondary growth demonstrated that cellulose synthase expression is confined exclusively to developing xylem cells undergoing secondary wall thickening. This cell-type-specific nature of *PrCelA* gene expression was also consistent with xylem-specific activity of cellulose synthase promoter (*PrCelAP*) based on heterologous promoter- β -glucuronidase (GUS) fusion analysis. Overall, the results provide several lines of evidence that cellulose synthase is the gene primarily responsible for cellulose biosynthesis during secondary wall formation in woody xylem of trees, such as aspen. Previous results by the inventors (Hu *et al.*, 1999) showed that cellulose and lignin are deposited in a compensatory fashion in wood. The discovery of a cellulose synthase in trees, such as aspen, permits the up-regulation of the protein to elevate cellulose production. Surprisingly, expression of *CelA* in trees suppressed lignin biosynthesis to further improve wood properties of trees.

Preparation of transgenic plants

The UDP-glucose binding sequence was subcloned into pBI121, which was used to prepare transgenic tobacco plants (Hu *et al.*, 1998). The expression of a heterologous UDP-glucose binding sequence resulted in a remarkable growth-accelerating effect. This was surprising because current knowledge of the function of plant cellulose synthases teaches that a UDP-glucose sequence must remain intact with other functional domains in *CelA*, e.g., the transmembrane domains, in order for cellulose synthase to initiate cellulose biosynthesis. The remarkable growth and tremendous increase in plant biomass observed in transgenic tobacco was due likely to an augmented deposition of

cellulose, indicating that the UDP-glucose domain alone is sufficient for genetic augmentation of cellulose biosynthesis in plants.

Genome organization and expression of a novel cellulose synthase

To confirm that the cDNA clone of Fig. 1 [SEQ ID NO: 1] was a cellulose synthase, genomic Southern blot analysis was performed under both high and low stringency conditions using the cDNA. Genomic DNA from aspen was digested with *Pst*I (lane P), *Hind*III (lane H) and *Eco*RI (lane E), and probed using a 1kb ³²P-labeled fragment from the 5' end of a cellulose synthase of Fig. 1. The Southern blot suggested the presence of a small family of cellulose synthase genes in aspen genome (Fig. 2, panels a and b). Repeated screening of the aspen xylem cDNA library with various plant *CelA* gene-related probes always resulted in the isolation of the same cellulose synthase cDNA clone. This suggested that the cellulose synthase cDNA cloned (Fig. 1) [SEQ ID NO: 1], represents the primary and most abundant cellulose synthase-encoding gene in developing xylem of trees, such as aspen, where active cellulose deposition takes place. It also indicates that manipulation of cellulose synthase gene expression can have a profound influence on cellulose biosynthesis in trees.

In situ hybridization

Northern blot analysis of total RNA from the internodes of aspen seedling stems (Fig. 2, panel c) using the labeled probe (as described above) revealed the near absence of cellulose synthase transcripts in tissues undergoing primary growth (internodes 1 to 4), and that the presence of cellulose synthase transcripts occurs during the secondary growth of stem tissues (internodes 5 to 11). However, weak northern signals in primary growth may only suggest that cellulose synthase gene expression is specific to xylem, of which there is little in primary growth tissue.

Xylogenesis in higher plants offers a unique model that involves sequential execution of cambium cell division, commitment to xylem cell differentiation, and culmination in xylem cell death (Fukuda, 1996). Although primary and secondary xylem cells originate from different types of cambia, namely procambium and inter/intrafascicular cambium, both exhibit conspicuous secondary wall development with massive cellulose and lignin deposition (Esau, 1965). To further investigate spatial and temporal cellulose synthase gene expression patterns at the cellular level, *in situ* hybridization was used to localize cellulose synthase mRNA along the developmental gradient defined by stem primary and secondary growth.

Localization of cellulose synthase gene transcripts (RNA) in stem at various growth stages was also observed. Fig. 3 shows transverse sections from 2nd, 4th and 6th internodes hybridized with digoxigenin (DIG)-labeled cellulose synthase antisense or sense (control) RNA probes, as described.

PtCesA transcripts were detected in young aspen stem sections by *in situ* hybridization with transcripts of highly variable 5' region of *PtCesA* cDNA (a 771 bp long fragment generated from *Pst*I and *Sac*I). This region was first subcloned in the plasmid vector, pGEM,-3Zf (+) (Promega) for the production of digoxigenin (DIG)-labeled transcripts using T7 (for antisense transcripts) and SP6 (for sense transcripts) RNA polymerase (DIG system: Boehringer Mannheim). Probes were subjected to mild alkaline hydrolysis by incubation in 100 mM NaHCO₃, pH 10.2 at 60 °C, which produced approximately 200 bp fragments.

Aspen young stems were prepared for sectioning by fixation in 4% (w/v) paraformaldehyde in 100 mM phosphate buffer (pH 7.0) at 4 °C overnight, dehydrated through an ethanol series on ice, and embedded in Paraplast medium (Sigma). Ten µm sections were mounted on Superfrost/plus (Fisher) slides at 42 °C overnight, dewaxed and then rehydrated through a descending ethanol series. The sections were incubated with proteinase K (10 µg/ml in 100 mM Tris-HCl, 50 mM EDTA, pH 7.5) for 30 min and were post-fixed with FAA. The sections were acetylated with 0.33% (v/v) acetic anhydride in 0.1 M triethanolamine-HCl (pH 8.0) prior to hybridization. The sections were then incubated in a hybridization mixture (approximately 2 µg/ml DIG-labeled probes, 50% (v/v) formamide, 2 X SSPE, 10% (w/v) dextran sulfate, 125 µg/ml tRNA, pH 7.5) at 45 °C for 12-16 hrs. Nonhybridized single-stranded RNA probe was removed by treatment with 20 µg/ml RNase A in TE buffer with 500 mM NaCl. The sections were washed at 50 °C. Hybridized DIG-labelled probe was detected on sections using anti-digoxigenin antiserum at a 1:1500 dilution, as described in the manufacturer's instruction (DIG system: Boehringer Mannheim). Sections were examined by Eclipse 400 light microscope (Nikon) and photographed.

During the primary growth stage (Fig. 3, panels a and b), strong expression of cellulose synthase was found localized exclusively to primary xylem (PX) cells. At this stage, young internodes are elongating, resulting in thickening of primary xylem cells through formation of secondary walls (Esau, 1968). The concurrence of shoot elongation with high expression of cellulose synthase strongly suggests the association of cellulose synthase protein with secondary cell wall cellulose synthesis. Later stages of primary growth (Fig. 3, panel b) are characterized by the appearance of an orderly alignment of primary xylem cells. Active cellulose biosynthesis accompanies cell elongation-induced wall thickening, as indicated by the strong expression of cellulose synthase in these primary xylem cells.

At the beginning of secondary growth in older internodes, it was observed that expression of cellulose synthase is also exclusively localized to xylem cells (Fig. 3, panel c). Instead of elongation in internodes distal to the meristematic activity, growth at this stage is mainly radial due to thickening in secondary cell walls of secondary xylem.

At the same time, expression of *PtCelA* gene becomes localized to the secondary developing xylem cells (SX in Fig. 3, panel c), which is again consistent with the idea that *PtCelA* encodes a secondary cell wall cellulose synthase. At this stage, secondary xylem cells cover the elongated and differentiated primary xylem cells in which *PtCelA* gene expression is no longer detectable (Fig. 3, panel c). These results demonstrate that expression of *PtCelA* gene is xylem-specific and the cellulose synthase of Fig. 1 [SEQ ID NO: 1] encodes a cellulose synthase associated with cellulose biosynthesis in secondary walls of xylem cells. To further confirm xylem-specific expression of cellulose synthase, a cellulose synthase gene promoter sequence was cloned and characterized for regulatory activities.

Characterization of expression regulated by cellulose synthase promoter

A 5' 1,200 bp cDNA fragment of a cellulose synthase of Fig. 1 [SEQ ID NO: 1] was used as a probe to screen an aspen genomic library for 5' regulatory sequences of a novel cellulose synthase gene, *PtCelA*. The library was constructed by cloning aspen genomic DNA fragments, generated from an *Sau3AI* partial-digest and sucrose gradient-selected, into the *Bam*HI site of a Lambda DASH II vector (Stratagene, La Jolla, CA). Five positive clones were obtained from about 150,000 pfu and Lambda DNA was purified. One clone having about a 20 kb DNA insert size was selected for restriction mapping and partial sequencing. This resulted in the identification of a 5' flanking region of *PtCelA* gene of approximately 1 kb. This genomic fragment, designated *PtCelAP* (Fig. 4) [SEQ ID NO: 3], contained about 800 bp of promoter sequence, 68 bp of 5' end untranslated region and 160 bp of coding sequence. To investigate regulation of tissue-specific cellulose synthase expression at the cellular level, promoter activity was analyzed in transgenic tobacco plants by histochemical staining of a GUS protein. A *PtCelAP-GUS* fusion binary vector was constructed in pBI121 with the 35S promoter replaced with *PtCelAP* [SEQ ID NO: 3] and introduced into tobacco (*Nicotiana tabacum*) as per Hu *et al.* (1998).

Eleven independent transgenic lines harboring a *CelAP-GUS* fusion were generated. Fig. 5 shows a histochemical analysis of *GUS* expression driven by a cellulose synthase promoter of the invention in transgenic tobacco plants. Transverse sections from the 3rd (panel a), 5th (panel b), 7th (panel c), and 8th (panels d and f) internodes were stained from GUS activity, and fluorescence microscopy was used to visualize expression under UV radiation.

GUS staining was detected exclusively in xylem tissue of stems, roots and petioles. In stems, strong GUS activity was found localized to xylem cells undergoing primary (Fig. 5, panel a) and secondary growth (Fig. 5 panels b-d and f). *GUS* expression was confined to xylem cells in the primary growth stage and became more localized in developing secondary xylem cells during secondary growth. An entire section from the

8th internode stained for GUS activity (Fig. 5, panel f). These results are consistent with the *in vivo* expression patterns of cellulose synthase in aspen stems. Lignin autofluorescence was visualized after UV radiation. Phloem fibers, which are also active in cellulose and lignin biosynthesis (Fig. 5, panels d and e), did not show GUS activity, suggesting that cellulose synthase gene expression is not associated with cellulose biosynthesis in cell types other than xylem. Examination of GUS activity in roots, stems, leaves, anthers and fruit also showed GUS expression in xylem tissue of all these organs suggesting that cellulose synthases of the invention are xylem-specific cellulose and expressed in all plant organs.

Characterization of promoter activity and cellular expression of a cellulose synthase of the invention from one particular source (aspen) indicated that expression produces a protein that encodes a secondary cell wall-specific cellulose synthase and is specifically compartmentalized in developing xylem cells. Characterization of the cellulose synthase gene promoter sequence not only confirms cell type-specific expression of cellulose synthase, but also provides a method for over-expressing cellulose synthase in a tissue-specific manner to augment cellulose production in xylem.

Expression of cellulose synthase under tension stress

As described earlier, a cellulose synthase promoter of the invention is involved in a novel gene regulatory phenomenon of cellulose synthase. To further characterize a cellulose synthase of the invention, *GUS* expression driven by an aspen cellulose synthase promoter (*PtCelAP*) was observed in transgenic tobacco plants without or under tension stress. The stress was induced by bending and affixing the plants to maintain the bent position (*e.g.*, tying) over a 40 hour period. Tangential and longitudinal sections were taken before bending, and 4 hrs, 20 hrs and 40 hrs after bending (panels a-d, respectively).

The cellulose synthase promoter-*GUS* fusion binary constructs showed exclusive xylem-specific expression of *GUS* without any tension stress (Fig. 6, panel a). However, under tension stress conditions endured by angiosperms in nature, the transgenic tobacco plants induced xylem and phloem-specific expression on the upper side of the stem within the first four hours of stress (Fig. 6, panel b).

This observation was surprising because during tension wood development fibers produce highly crystalline cellulose in order to provide essential mechanical strength to a bending stem. The present observation was the first showing of transcriptional up-regulation of a cellulose synthase, mediated through a cellulose synthase promoter that is directly responsible for development of highly crystalline cellulose in trees. Furthermore, after 20 hrs of tension stress, both xylem and phloem exhibited *GUS* expression, but only on the upper side of the stem that was under tensile stress, *i.e.*, *GUS* expression on the lower side was inhibited (Fig. 6, panel c). With extended stress (up to 40 hrs), *GUS*

expression was restricted to only one small region on the upper side of the stem where maximum tension stress was present (Fig. 6, panel d). Based on the observation of GUS signal in woody cells upon tension stress and the absence of GUS under compression or no stress, it was concluded that a cellulose synthase promoter of the invention has mechanical stress responsive elements (MSREs) that turn cellulose synthase genes on and off depending on the presence and type of stress to the stem.

The results indicate that positive MSREs exist in a cellulose synthase promoter of the invention to bind transcription factors in response to tension stress for regulating the expression of cellulose synthase and increasing biosynthesis of higher crystalline cellulose. This is evident based on the expression of GUS in xylem and phloem tissue at the upper side of the stem subjected to tension stress, but not when tissue on the lower side was subjected to compression or no stress. Furthermore, the tissue at the lower side of the stem, which was subjected to compression stress, showed no GUS expression, i.e., expression was turned off. This indicated the presence of negative MSREs, which bind transcription factors to turn off expression of cellulose synthase at the lower side of the stem. Negative MSREs likely suppress development of highly crystalline cellulose in normal wood.

These results provide a mechanism for genetically engineering synthesis of highly crystalline cellulose in juvenile wood for enhancing strength properties, and for synthesizing a higher percentage of cellulose in reaction wood. The positive MSREs and their cognate transcription factors are important in the synthesis of highly crystalline cellulose of high tensile strength, as are the negative MSREs and inhibition of cognate transcription factors thereto. The present invention thus provides a starting point for cloning cDNAs for the transcription factors that bind to positive and negative MSREs according to methods known in the art. Constitutive expression of cDNAs for positive MSRE transcription factors allows the continuous production of highly crystalline cellulose in transgenic trees, while expression of antisense cDNAs for negative MSRE transcription factors inhibits those transcription factors so that cellulose synthase cannot turn off. This combination will assure continuous production of highly crystalline cellulose in trees.

Genetic engineering of cellulose synthase in transgenic plants

As discussed above, the nucleotide sequence of a cellulose synthase of the invention, e.g., *PtCelA* cDNA from aspen, shows significant homology with other polynucleotides encoding cellulose synthase proteins that have been suggested as authentic cellulose synthase clones. To further characterize the activity of a cellulose synthase, four constructs were prepared in a PBI121 plasmid.

1) A constitutive plant promoter Cauliflower mosaic Virus 35S was operatively linked to *PtCelA* (35SP-PtCelA-s) and overexpressed in transgenic plants.

This causes excess production of cellulose, resulting in a reduction in lignin content. Tobacco and aspen have been transformed with this construct.

2) Cauliflower mosaic Virus 35S was operatively linked to antisense RNA from *PtCela* (35S-*PtCela*-a) and constitutively expressed to reduce production of cellulose and increase lignin content in transgenic plants. This negative control construct may not result in healthy plants since cellulose is essential for plant growth and development. Aspen plants have been transformed with this construct.

3) Aspen 4CL-1 promoter (Hu *et al.*, 1998) was operatively linked to *PtCela* (Pt4CLP-*PtCela*) (the 35S promoter of pBI121 was removed in this construct) and expressed in a tissue-specific manner in developing secondary xylem of transgenic aspen. This expression augments the native cellulose production and reduces lignin content of angiosperm tissues. Tobacco and aspen have transformed with this construct.

4) The cytoplasmic domain of *PtCela* which contains three conserved regions thought to be involved in UDP-glucose binding during cellulose biosynthesis, was linked to a 35S promoter to produce binary constructs (35S-*PtCela* UDP-glucose). Expression by this promoter permits constitutive expression of a UDP glucose binding domain of *PtCela* in transgenic plants. Tobacco and aspen have been transformed with this construct.

35S-*GUS* constructs (pBI121, ClonTech, CA) were used as controls for each experiment with the constructs. Transgenic tobacco plants were transformed with the constructs. The following table shows the general growth measurements of the T0 tobacco plants. Plants carrying a *PtCela* construct grew much faster than control plants carrying a pBI121 (control) construct. In comparing developmental 4CL and constitutive 35S promoter control of *PtCela* expression, the 35S was more effective, permitting faster growth of transgenic tobacco plants. The fastest growth was seen in transgenic plants carrying a 35S promoter driven UDP-G domain from *PtCela*.

It is noted that T0 generation plants can have carry over effects from their tissue culture treatments. Therefore, seeds were collected for testing this growth phenomenon in T1 generations. The transgenic tobacco plants were analyzed for presence of the transferred genes and all tested positive for the respective gene constructs.

TABLE
Transgenic tobacco plant measurements
after transfer in soil for about 1.5 months (N = 2)

Construct	Height	Diameter	Internode length	No. of leaves	Longest leaf
35S-GUS	17	0.5	1	11	17
35S-PtCelA	77	1.0	6	13	37
35S-UDPG	83	1.0	6	13	37
4CLP-PtCelA	41	0.8	5	10	29

Note: All values were measured in centimeters, excluding number of leaves.

5

It will be appreciated by persons of ordinary skill in the art that the examples and preferred embodiments herein are illustrative, and that the invention may be practiced in a variety of embodiments which share the same inventive concept.

10

BIBLIOGRAPHY

- Hu et al.*, 1999, Nature Biotechnology, In Press
- Whetten et al.*, 1998, Ann Rev Pl Physiol Pl Mol Biol, 49: 585-609
- 5 *Arioli et al.*, 1998, Science, 279: 717-720
- Wu et al.*, 1998, Pl Physiol, 117: 1125
- Hu et al.*, 1998, PNAS, 95: 5407-5412
- Joshi et al.*, 1997, PMB, 35: 993-1001
- Fukuda*, 1996, Ann Rev Pl Physiol Pl Mol Biol, 47: 299-325
- 10 *Pear et al.*, 1996, PNAS, 93: 12637-12642
- Haigler and Blanton*, 1996, PNAS, 93: 12082-12085
- Ge and Chiang*, 1996, Pl Physiol, 112: 861
- Brown et al.*, 1996, Trends Pl Sci., 1: 149-156
- Delmer and Amor*, 1995, Pl Cell, 7: 987-1000
- 15 *Hoffman and Stoffel*, 1993, Biol Chem, Hoppe-Seyler 374: 166
- Joshi*, 1987, NAR, 15: 6643-6653
- Joshi*, 1987, NAR, 15: 9627-9640
- Timmell*, 1986, Compression Wood in Gymnopserms, Springer Verlag
- Esau*, 1967, Plant Anatomy, Wiley and sons, NY
- 20 *Higuchi*, 1997, Biochemistry and Molecular Biology of Wood, Springer Verlag

WHAT IS CLAIMED IS:

1. A polynucleotide which is (a) a polynucleotide having a nucleotide sequence of SEQ ID NO:1; (b) a polynucleotide having a nucleotide sequence of SEQ ID NO: 4; or (c) a polynucleotide fragment of (a) or (b) encoding a functional domain of a cellulose synthase.
2. The polynucleotide of claim 1 wherein the polynucleotide is operably linked to a polynucleotide of SEQ ID NO: 3, or a functional fragment thereof.
3. A vector comprising a polynucleotide of claim 1.
4. A transgenic plant comprising a polynucleotide of claim 1.
5. A cellulose synthase promoter, or a functional fragment thereof which binds a transcription factor in a plant cell.
6. A vector comprising a promoter or a fragment of claim 5.
7. A transgenic plant comprising a promoter or a fragment of claim 5.
8. A polypeptide having an amino acid sequence of SEQ ID NO: 2, an amino acid sequence of SEQ ID NO: 5 or an amino acid sequence sequence which a functional domain of cellulose synthase.
9. A method of altering the growth of a plant, comprising expressing in cells of the plant an exogenous polynucleotide encoding a cellulose synthase wherein the polynucleotide is expressed in an amount effective to alter the growth of the plant.
10. A method according to claim 9, wherein the polynucleotide comprises a homologous polynucleotide.
11. A method according to claim 9, wherein the polynucleotide comprises a heterologous polynucleotide.
12. A method according to claim 9, wherein the polynucleotide is in a sense orientation.

13. A method according to claim 9, wherein the polynucleotide is in an antisense orientation.
14. A method according to claim 9, wherein a plant promoter, or a transcription
5 factor binding domain thereof, is operatively linked to the polynucleotide.
15. A method according to claim 14, wherein the promoter is selected from constitutive promoters, tissue-specific promoters and developmental-specific plant
10 promoters.
16. A method according to claim 15, wherein the promoter is Cauliflower Mosaic Virus 35S, 4CL, cellulose synthase promoter, *PtCelAP* or terminal flower promoter.
- 15 17. A polynucleotide encoding a UDP-glucose binding domain of a cellulose synthase.
18. A polypeptide comprising a UDP-glucose catalytic subunit of cellulose synthase wherein the UDP-glucose catalytic subunit catalyzes the biosynthesis of
20 cellulose.
19. A method of altering the growth of a plant, comprising incorporating into a plant genome a polynucleotide encoding a UDP-glucose catalytic subunit wherein expression of the polynucleotide alters the growth of the plant.
- 25 20. A method according to claim 19, wherein the polynucleotide comprises a homologous polynucleotide.
21. A method according to claim 19, wherein the polynucleotide comprises a
30 heterologous polynucleotide.
22. A method according to claim 19, wherein the polynucleotide is in a sense orientation.
- 35 23. A method according to claim 19, wherein the polynucleotide is in a antisense orientation.

24. A method of altering the cellulose content in a plant comprising expressing an exogenous polynucleotide encoding a UDP-glucose binding domain in a plant genome to alter the cellulose content of the plant.

5 25. A transgenic plant having an increased ratio of cellulose to lignin in cells of the plant comprising an exogenous polynucleotide encoding a cellulose synthase operably linked to a promoter so that the polynucleotide is expressed in an amount effective to increase the cellulose content of the plant.

10 26. A transgenic plant having a decreased ratio of lignin to cellulose, the plant comprising an exogenous polynucleotide encoding a cellulose synthase.

27. A method of altering expression of a cellulose synthase in a plant cell comprising delivering into the cell a vector comprising a polynucleotide encoding a cellulose synthase.
15

28. The method according to claim 27, wherein the polynucleotide comprises a homologous polynucleotide.

20 29. The method according to claim 27, wherein the polynucleotide comprises a heterologous polynucleotide.

30. The method according to claim 27, wherein the polynucleotide is in a sense orientation.

25 31. The method according to claim 27, wherein the polynucleotide is in an antisense orientation.

30 32. A method of causing stress-induced gene expression in a plant cell comprising delivering into the cell a vector comprising a cellulose synthase promoter operatively linked to a gene, wherein the gene is expressed upon a mechanical stress to the plant.

33. A method of determining a positive mechanical stress responsive element (MSRE) in a cellulose synthase promoter comprising:

(i) introducing into a plant a cellulose synthase promoter that has a portion deleted, the cellulose synthase promoter operatively linked to a polynucleotide encoding a reporter, and

(ii) detecting a decrease in the amount of reporter in the plant after inducing a stress to the plant.

34. A method of determining a negative MSRE in a cellulose synthase promoter comprising:

(i) introducing into a plant a cellulose synthase promoter that has a portion deleted, the cellulose synthase promoter operatively linked to a reporter gene, and

(ii) detecting an increase in the amount of reporter in the plant after inducing a stress to the plant.

35. A method of expressing cellulose synthase in a plant in a tissue-specific manner comprising transforming the plant with a tissue-specific promoter operatively linked to a polynucleotide encoding a cellulose synthase.

36. A method of increasing expression of a cellulose synthase in a plant comprising delivering into the plant a cDNA encoding a protein that binds to a positive MSRE of a cellulose synthase promoter wherein the binding to the positive MSRE results in expression of a cellulose synthase, resulting in increased expression of cellulose in the plant.

37. A method of reducing expression of a cellulose synthase in a plant comprising delivering into the plant a cDNA in an antisense orientation, the cDNA in a sense orientation encoding protein that binds to a positive MSRE and results in expression of a cellulose synthase.

38. A method of increasing cellulose biosynthesis in a plant comprising delivering into the plant a cDNA encoding a protein that binds to a positive MSRE of a cellulose synthase promoter, wherein binding of the protein to the positive MSRE results in expression of a cellulose synthase.

39. A method of reducing cellulose biosynthesis in a plant comprising delivering into the plant a cDNA in an antisense orientation, the cDNA in a sense orientation encoding protein that binds to a positive MSRE of a cellulose synthase promoter.

40. A transgenic plant containing a polynucleotide comprising a promoter and encoding a cellulose synthase, the polynucleotide expressed such that the growth of the plant is altered relative to a similar control plant that does not contain the polynucleotide.

5 41. A vector comprising a promoter functional in a plant cell, and a coding sequence for cellulose synthase, the coding sequence operably linked to the promoter, the promoter having a nucleotide sequence encoding a positive MSRE of cellulose synthase.

42. A method of altering a characteristic of a plant comprising genetically
10 upregulating cellulose synthase in the plant, wherein the characteristic is accelerated growth, increased cellulose content or decreased lignin content.

43. The method of claim 42 wherein the plant is genetically upregulated
15 through incorporation into the genome of the plant a cDNA having a nucleotide sequence encoding a cellulose synthase.

44. A method of regulating cellulose synthase expression in a plant comprising
delivery in a plant (a) a cDNA encoding a polypeptide which is a positive MSRE of a
cellulose synthase promoter; or (b) a cDNA in an anti sense orientation of the cDNA of
20 (a), in amount and under conditions effective to allow at least a portion of the plant's cells to take up the cDNA.

45. A method of altering cellulose content in a plant comprising:
(a) delivery into cells of the plant an expression cassette comprising a
25 cDNA encoding a cellulose synthase operably linked to a promoter functional in a plant cell; and

(b) expressing the cDNA in the cells of the plant in an amount effective to alter the cellulose content in the cells of the plant.

30 46. A DNA encoding a protein having cellulose synthase activity and comprising the amino acid sequence in SEQ ID NO:2, SEQ ID NO:5 or an amino acid sequence which is a functional domain of cellulose synthase.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 November 2000 (30.11.2000)

PCT

(10) International Publication Number
WO 00/71670 A3

- (51) International Patent Classification: C12N 15/05, 15/09, 15/29, 15/52, 15/82, A01H 5/00
- (21) International Application Number: PCT/US00/13637
- (22) International Filing Date: 18 May 2000 (18.05.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/135,280 21 May 1999 (21.05.1999) US
- (71) Applicant (for all designated States except US): BOARD OF CONTROL OF MICHIGAN TECHNOLOGICAL UNIVERSITY [US/US]; 1400 Townsend Drive, Houghton, MI 49931 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): CHIANG, Vincent, L. [US/US]; 1104 Birch Street, Hancock, MI 49930 (US), WU, Luguang [CN/AU]; 123 Akana Street, Kenmore, QLD 4069 (AU). CHANDRASHEKHAR, Joshi, P. [IN/US]; 713 Cedar Bluff Drive, Houghton, MI 49931 (US). CARRAWAY, Daniel, T. [US/US]; 1910 Rich Street, Bainbridge, GA 31717 (US).
- (74) Agent: GEMIGNANI, Joseph, A.; Michael Best & Friedrich LLP, 100 East Wisconsin Avenue, Milwaukee, WI 53202 (US).
- (81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— With international search report
— Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- (88) Date of publication of the international search report: 18 January 2001
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette

(54) Title: METHOD FOR ENHANCING CELLULOSE AND MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS

(57) Abstract: This invention relates to polynucleotide molecules encoding cellulose synthase, promoters of cellulose synthase and cellulose synthase polypeptides, methods for genetically altering cellulose and lignin biosynthesis, and methods for improving strength properties of juvenile wood and fiber in trees. The invention further relates to methods for identifying regulatory elements in a cellulose synthase promoter and transcription factors that bind to such regulatory elements, and to methods for augmenting expression of polynucleotides operably linked to a cellulose synthase promoter.

WO 00/71670 A3

1/10

Fig. 1: DNA and predicted protein sequence of PtCela cDNA

1	GTCGACCCACGCGTCCGCTTGAAGAAATATGAAGTTGTAAGAGCTGGTAAAGTGGTAA	60
61	TAAGCAAGATGATGGAATCTGGGGCTTCTATATGCCATACCTGTGGTGAACAGGTGGGGC M M E S G A P I C H T C G E Q V G H	120
121	ATGATGCAAAATGGGAGCTATTTGTCTCTGCCATGAGTGTAGCTATCCCATGTGCAAGT D A N G E L F V A C H E C S Y P M C K S	180
181	CTTGTTTCGAGTTTGAAGTCAATGAGGGCGGAAAGTTTGGCTTGGCGTGTGGCTCGCCAT C G F E F E I N E G R K V C L R C G S P Y	240
241	ATGATGAGAAGTCTGCTGGATGATGTAGAAAAGAAGGGGTCTGGCAATCAATCCACAATGG D E N L L D D V E K K G S G N Q S T M A	300
301	CATCTCACCTCAACGATTTCTCAGGATGTCGGAATCCATGCTAGACATTCAGTAGTGTGT S H L N D S Q D V G I H A R H I S S V S	360
361	CCACTGTGGATAGTGAAATGAATGATGAATATGGGAATCCAATTGGAAGAATCGGGTGA T V D S E M N D E Y G N P I W K N R V K	420
421	AGAGCTGTAAGGATAAAGAGAACAGAGAGAAAAGAGAGTCCCTAAGGCTGAAACTGAAC S C K D K E N K K K K R S P K A E T E P	480
481	CAGCTCAAGTTCCTACAGAACAGCAGATGGAAGAGAAACCGTCTGCAGAGGCTTCGGAGC A Q V P T E Q Q M E E K P S A E A S E P	540
541	CGCTTTCAATTGTTTATCCAATTCACGCAACAAGCTCACACCATACAGAGCAGTGATCA L S I V Y P I P R N K L T P Y R A V I I	600
601	TTATGCGACTGGTCAATCTGGGCCCTCTTCTCCACTTCAGAATAACAAATCTGTGCGATA M R L V I L G L F F H F R I T N P V D S	660
661	GTGCTTTTGCCCTGTGGCTTACTTCTGTCTATGTGAGATCTGGTTTGCATTTTCTTGGG A F G L W L T S V I C E I W F A F S W V	720
721	TGTTGGATCAGTTCCCAAGTGAATCCTGTCAATAGAGAAACGTATATCGAAAGGCTGT L D Q F P K W N P V N R E T Y I E R L S	780
781	CGGCAAGGTATGAAGAGAGGGGTGAGCTTCTCAGCTTGCTGGTGTGGATTTTTCGTGA A R Y E R E G E P S Q L A G V D F F V S	840
841	GTACTGTTGATCCGCTGAAGGAACCGCCATTGTACTGCCAATCAGTCTCTTCCATTC T V D P L K E P P L I T A N T V L S I L	900
901	TGCTGTGGAAGTATCCCGTCGATAAAGTCTCTGCTACGTGTCTGATGATGGTGCAGCTA A V D Y P V D K V S C Y V S D D G A A M	960
961	TGCTTTCATTGAATCTCTTGTAGAACAGCTGAGTTTGAAGGAAGTGGGTTCCGTTCCT L S F E S L V E T A E F A R K W V P F C	1020
1021	GCAAAAAATCTCAATTGAACCAAGAGCACCGGAGTTTACTTCTCACAGAAAATTGATT K K F S I E P R A P E F Y F S Q K I D Y	1080
1081	ACTTGAAAGACAAGGTTCAACCTTCTTTCGTGAAAGAACGTAGAGCAATGAAAAGGGATT L K D K V Q P S F V K E R R A M K R D Y	1140
1141	ATGAAGAGTACAAAGTCCGAGTTAATGCCCTGGTAGCAAGGCTCAGAAAACACCTGAAG E E Y K V R V N A L V A K A Q K T P E E	1200

2/10

1201	AAGGATGGACATATGCAAGATGGAACACCTTGGCCTGGGAATAACACACGTGATCACCCCTG G W T M Q D G T P W P G N N T R D H P G	1260
1261	GGCATGATTCAGGTCTCTCTTGGGAAATACTGGGAGCTCGTGACATTGAAGGAAATGAAC H D S G L P W E I L G A R D I E G N E L	1320
1321	TACCTCGTCTAGTATATGTCTCCAGGGAGAAGAGACCTGGCTACCAGCACCACAAAAAGG P R L V Y V S R E K R P G Y Q H H K K A	1380
1381	CTGGTGCAGAAAATGCTCTGGTGAGAGTGTCTGCAGTACTCACAAATGCTCCCTACATCC G A E N A L V R V S A V L T N A P Y I L	1440
1441	TCAATGTTGATTGTGATCACTATGTAAACAATAGCAAGGCTGTTCGAGAGGCAATGTGCA N V D H Y V N N S K A V R E A M C I	1500
1501	TCCTGATGGACCCACAAGTAGGTGCGAGATGTATGCTATGTGCAGTTCCCTCAGAGGTTTG L M D P A Q G V G R D V C Y V Q F P Q R F D	1560
1561	ATGGCATAGATAAAGATGTACGCTACGCCAATCGTAACGTAGTTTCTTTGTGTTTAAACA G I D K S D R Y A N R N V V F F D V N M	1620
1621	TGAAAGGGTTGGATTGGCATTCAAGGACCAGTATACGTAGGAACCTGGTCTGTGTTTCAACA G K L D W I Q G P V Y V G T G C T V F N R	1680
1681	GGCAAGCACTTTACGGCTACGGGCTCCCTTCTATGCCAGCTTACGCAAGAGAAAGGATT Q A L Y G Y G P P S M P S L R K R K D S	1740
1741	CTTCACTCCCTGCTTCTCATGTTGCTGCCCTCAAGAAGAAGCCCTGCTCAAGATCCAGCTG S S C F G S C C P S K K K P A Q D V F A E	1800
1801	AGGTATACAGAGATGCAAAAAGAGAGGATCTCAATGCTGCCATATTAACTCTTACAGAGA V Y R D A K R E D L N A A I F N L T E I	1860
1861	TTGATAATTATGACGAGCATGAAAGGTCAATGCTGATCTCCAGTTGAGCTTTGAGAAAA D N Y D E H E R S M L I S Q L S F E K T	1920
1921	CTTTTGGCTTATCTTCTGTCTTCATTGAGTCTACACTAATGGAGAATGGAGGAGTACCCG F G L S S V F I E S T L M E N G G V P E	1980
1981	AGTCTGCCAATCAACCACCTTCAATCAAGGAAGCGATTCAAGTCATCGGCTGTGGCTATG S A N S P P F I K E A I Q V I G C G Y E	2040
2041	AAGAGAAGACTGAATGGGGGAAAACAGATTGGTTGGATATATGGGTCAAGTCATGAGGATA E K T E W G K Q I G W I Y G S V T E D I	2100
2101	TCTTAAGTGGCTTCAAGATGCACTGCCGAGGATGGAGATCAATTACTGCATGCCCGTAA L S G F K M H C R G W R S I Y C M P V R	2160
2161	GGCTGCATTCAAAGGATCTGCACCCATCAACCTGTCTGATAGATTGCACAGGTCCTCC P A F K S A P I N L S D R L A H Q V L R	2220
2221	GATGGGCTCTTGGTTCTGTGGAAATTTCTTTAGCAGACACTGTCCCTCTGGTACGGGT W A L G S V E I F F S R H C P L W Y G F	2280
2281	TTGGAGAGGGCGCTTTAAATGGCTCCAAAGGCTTGCCTATATAAAACACCATTTGTGTAC G G R L K W L Q R L A Y I N T I V Y P	2340
2341	CATTACATCCCTCCCTCTCATTGCCATTGACACAATCCCTGCAGTTTGCTGCTCACC F T S L P L I A Y C T I P A V C L L T G	2400
2401	GAAATTCATCATCAACAGGCTCTCAACCTGGCAAGCATGCTGTTCTTGGCCCTTTTA K F I I P T L S N L A S M L Q L F I	2460
2461	TCTCCATCATTGTAAGTGGGTGCTTGAGCTAAGATGGAGCGGTGTCAGCATTTGAAGATT	2520

3/10

S I I V T A V L E L R W S G V S I E D L
2521 TATGGCGTAATGAACAATCTGGGTGATCGGAGGTGTTTCAGCCCATCTCTTTGCGGTCT 2580
W R N E Q F W V I G G V S A H L F A V F
2581 TCCAGGGATTCTTAAAAATGTTGGCTGGCATCGATACGAACTTCACTGTACAGCAAAAAG 2640
Q G F L K M L A G I D T N F T V T A K A
2641 CAGCCGAAGATGCAGAATTTGGGGAGCTATATATGGTCAAGTGGACAACACTTTTGATT 2700
A E D A E F G E L Y M V K W T T L L I P
2701 CTCCAACCACTTCTCATATCAATATGTCGGGTGTGTGCTGGATTCTCTGATGCACTCA 2760
P T T L L I I N M S G C A G F S D A L N
2761 ACAAGGATATGAAGCATGGGGGCTCTCTTTGGCAAGGTGTTCTTTGCTTTCTGGGTGA 2820
K G Y E A W G P L F G K V F F A F W V I
2821 TTCTTCATCTCTATCCATTCCTTAAAGGTCTAATGGGTCGCCAAAACCTAACACCAACCA 2880
L H L Y P F L K G L M G R Q N L T P T I
2881 TTGTTGTTCTCTGGTCAGTGTCTGTTGGCTCTGTCTTCTCTCTCGTTTGGGTCAAGATCA 2940
V V L W S V L L A S V F S L V W V K I N
2941 ATCCATTGCTTAAACAAAGTTGATAACACCTTGGTTGCGGAGACCTGCATTTCCATTGATT 3000
P F V N K V D N T L V A E T C I S I D C
3001 GCTGAGCTACCTCCAATAAGTCTCTCCCACTATTTTGGGGTTACAAAACCTTTGGGAATT 3060
*
3061 GGAATATGATCTCGTTGTAGTTTCCCTCAAGAAAGCACATATCGCTGTCAGTATTTAAA 3120
3121 TGAAC TGCAAGATGATTGTCTCTATGAAGTTTGAACAGTTTGAATGATATTATGTTA 3180
3181 AAATACAGGTTTTGATTGTGTGAAAAAAGAAAAA 3232

4/10

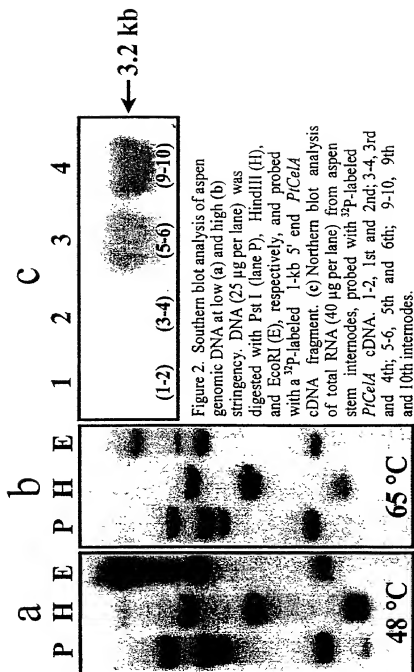


FIG. 2

5/10

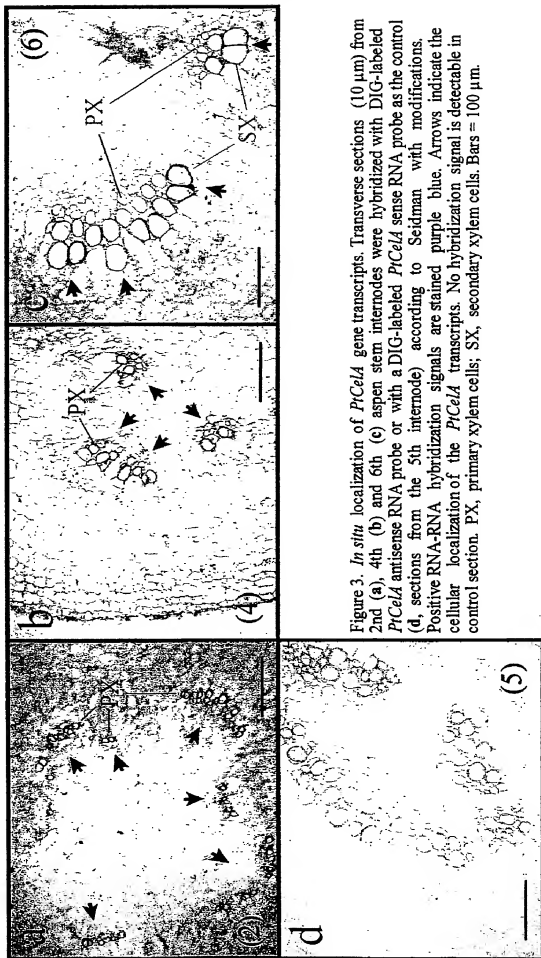


Figure 3. *In situ* localization of *Picea* gene transcripts. Transverse sections (10 μm) from 2nd (a), 4th (b) and 6th (c) aspen stem internodes were hybridized with DIG-labeled *Picea* antisense RNA probe or with a DIG-labeled *Picea* sense RNA probe as the control (d, sections from the 5th internode) according to Seidman with modifications. Positive RNA-RNA hybridization signals are stained purple blue. Arrows indicate the cellular localization of the *Picea* transcripts. No hybridization signal is detectable in control section. PX, primary xylem cells; SX, secondary xylem cells. Bars = 100 μm.

FIG. 3

Fig. 4: DNA sequence of PtCelAP, the 5' flanking region of PtCelA coding sequence

1	GAATTCGCCCCTTTTGAATTCAGGAGACGATAGTTTCCGGTTCGTTGAATGGCTTTGTTC	60
61	CTTCTGGCTAGCAATTTGCAAAAGAAGTTACAAAACAAATGCATATTATGTAAATTTAA	120
121	CAAGAGATGGTTCATGGTCACTTATTTATGCCCATCATTTGTTCTGGGGTTACTCTTT	180
181	ATAGTCTGATTCGAAGTTGCAAACTGCCGTTTCTGGTATTGCAATTATGTAGCCATAAAC	240
241	TGTTAATCCGTAGCTATTAGCGGACCAACAACCAGATATACGGGATCAGCGTCGTAAAA	300
301	GAGATCCTCATTTCTACGTTTCTTTCTAATTTTCCGTTTTCAGTGAGAGAATTACCTGAT	360
361	ACATTGACATGATGATTGATGATTATGGGAACCATTCGGATGTTAGACACGAGACCATCT	420
421	GGATCCTGCCAGTTTTCAGTTCACATGGCATCTCAGCCCAAGATCATGTGTTTATACGCC	480
481	TAATGACTTGTATTGAAAGTTTGGTAAGTTGAAGATGTGCTCTGCCCAACAGAAACCTTC	540
541	CTTAAATTTCCAGCAATCTTTCAAACCTTGGCCTTACACCCCGAAAAATAGACGTGCTTCT	600
601	ACTTGGGTTCTTGGAAACCATGCACCAACCGCCATACCCACCAACCCACCCCTCAAC	660
661	CTTCTCTTCGCCATTACAAAAATGTCAGTACCACCCCTCTGAAAGACACCAACACCCCTA	720
721	GCTTTGGTTAGGGTATTTGATATAAAAACAAGGCCAAAAACAAAGATTGGAAGGAAGCAG	780
781	AGGAAGACCCCTCTTGAAAGAATTGAAGTTGTAAAGAGCTGGTAAAGTGGTAATAAGCAAG	840
841	ATGATGGAATCTGGGGCTCCTATATGCCATACCTGTGGTGAACAGGTGGGGCATGATGCA	900
	M M E S G A P I C H T C G E Q V G H D A	
901	AATGGGGAGCTATTTGTGGCTTGCCATGAGTGTAGCTATCCCATGTGCAAGTCTTGTTC	960
	N G E L F V A C H E C S Y P M C K S C F	
961	GAGTTGAAATCAAGAGGGCCGAAAGTTTGCTTGCGGTGTGGCTCGAG	1010
	E F E I K E G R K V C L R C G S	

7/10

FIG. 5

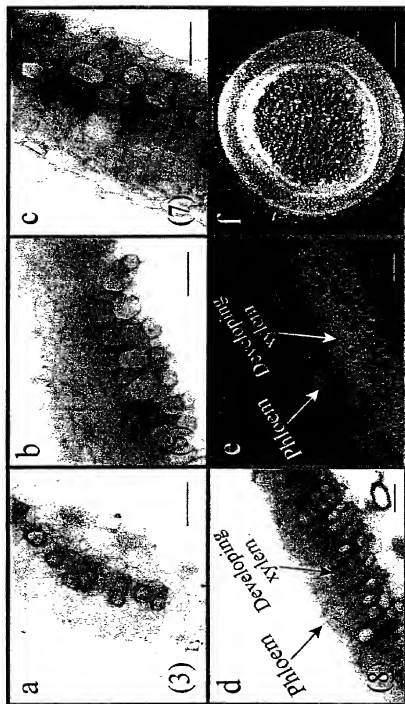


Figure 5. Histochemical analysis of transgenic tobacco for *GUS* gene expression driven by aspen *PtCel4* gene promoter. Stem transverse sections from the 3rd (a), 5th (b), 7th (c), and 8th (d and f) internodes were stained for *GUS* activity. Fluorescence microscopy (e) showing the identical section as in (d). Lignin autofluorescence was visualized after UV irradiation. An entire section from the 8th internode stained for *GUS* activity is shown in (f). Bars = 100 μ m in a to e; bar = 1.5 mm in f.

8/10

FIG. 6

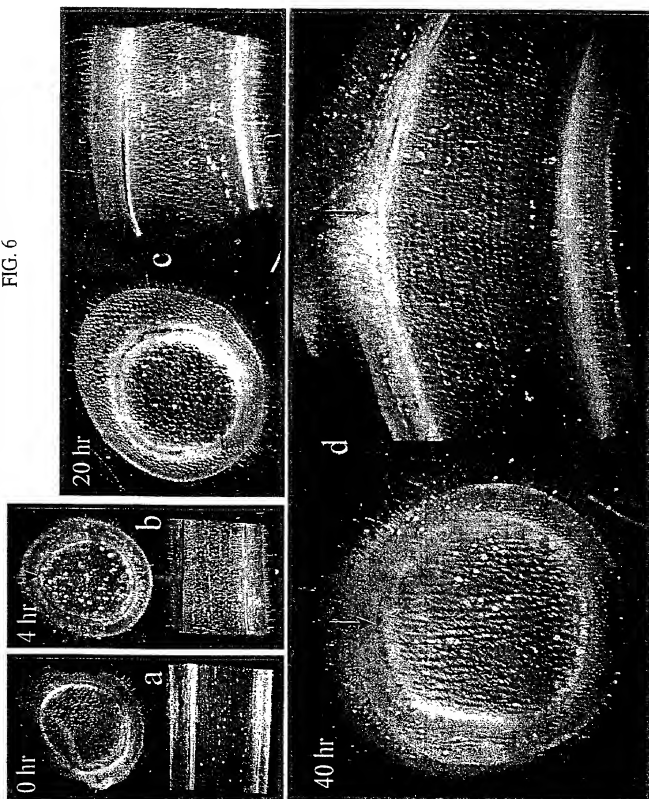


Figure 6. GUS gene expression driven by aspen P_{Ca} gene promoter in transgenic tobacco plants under tension stress. Tension stress was induced by bending the transgenic plants. Tangential and longitudinal sections were harvested before bending (a) and 4 (b), 20 (c) and 40 (d) hours after bending and stained for GUS expression. Arrows indicate the bend sites.

9/10

FIG. 7

Arabidopsis thaliana cellulose synthase mRNA SEQ ID NO: 4

1	gcggccgcg	ttaatccgcg	gtttccacaa	caggaatgag	ttgtctctca	ttaatgccga
61	tgagaatgcc	cgaaataagat	cagtcccaaga	gtcgagtggga	cgagacatgtc	aaatctgcag
121	agatgagatc	gaacttgactg	ttgatgggaga	accgtttgtg	gcatgttaacg	aatgtgcatt
181	ccctgtgtgt	agaccttgct	atgagtacga	aagacgagaa	ggcaatcaag	cttctgccca
241	gtgcaaaacc	cgtttcaaac	gtcttaaagg	aagtccaaga	gttgaaaggtg	atgaagagga
301	agatgacatt	gatgatttag	acaatgagtt	tgagtatgga	aataatggga	ttggatttga
361	tcagggtttct	gaaggatagt	caatctctcg	tcgcaactcc	ggtttcccac	aatctgattt
421	ggatttcagct	ccacctggct	ctcagattcc	attgctgact	tcaggcgacg	aggacgttga
481	gatttctctct	gatagacatg	ctcttattgt	tcctctcttca	ctcgggtggtc	atggcaatag
541	agttcatcct	gttctctctt	ctgacccgac	cggtgcttga	catcgaaagg	tgatggtagc
601	tcgaaagat	cttgcgggtt	atggttatgg	aagtgctcgt	tggaaagatc	ggatggagga
661	atggaagaga	aagcagaatg	agaaacttca	ggttgttagg	catgaaggag	atcctgattt
721	tgaagatggg	gatgatcgct	attttccaat	gatggatgag	ggaaggcagc	catgtctcat
781	gaagataacca	atcaaattcga	gcaagataaaa	tccttaccgg	atgtttaattg	tgctacgtgt
841	ttgtgattctt	ggatctctct	ttcactaccg	tattcttcac	ccggtcaaa	atgcataatg
901	tttgtggtct	atttctgtta	tatgtgagat	atggtttgct	gtttcatggg	ttcttgatca
961	gttccctaaa	tggtaaccta	tcgagcgaga	aacgtacttg	gaccgactct	cattaagata
1021	tggaagaaga	gggaaccgct	cgggactatc	ccctgtggat	gtattttgta	gtacagtggga
1081	tccattgaaa	gagctccgcg	ttattactgc	aaatactgtc	ttgtctatct	ttcttgatca
1141	ttatctctgt	gataaggttg	cttgttacct	atctgatgat	gggtgcgtca	tgcttacttt
1201	cgaaggctct	ctgagacgcg	ctgaattcgc	aaggaataatg	gttctcttct	cgacgaataa
1261	ttgtatttgc	ctctgctctc	ccgaatggta	tttctgccat	aaaatggact	acttgaagaa
1321	taaaagttcat	ccccctattg	ttagggaagg	gcgagccatg	aagagaagat	atgaagaatt
1381	caaaagttaag	atcaatgctt	tagtagcaac	agcagagaaa	gtgcctgagg	atgggtggac
1441	tatgcaagac	ggatcacctt	ggcccggtaa	tagtgtgcga	gcattctctg	gcattgattca
1501	ggctcttctt	ggagtgagcg	gtgttcctga	tgtcgaaaac	aacgagttgc	ctcgatttag
1561	ttacgtttct	cgtgagaaga	gaccgggatt	tgatcaccat	aagaaggctg	gcgcatgagt
1621	ttcccttgata	cgagttctct	gggttctatc	aaatgctcct	tacctcttga	atgtcgattg
1681	tgatcactac	atcaacaata	gcaaaagctc	tagagaagca	atgtgtttca	tgatggatcc
1741	tcagtcaggga	aagaaaaatct	gttatgttca	gttccctcaa	aggttctgat	ggatgtgat
1801	gcacgatcgga	tactcaaatc	gcaatgttgt	gttctttgat	atcaattatga	aaggtttgga
1861	tgggtctaca	gggctctaat	acgtcggtag	aggtgtgtct	ttcaggaggg	aagcgcttta
1921	cgaatttgat	gcaaccgaaga	agaagagagg	cccacgtaa	acatgcaatt	gtgggcaaa
1981	atgggtctct	ctatgttttg	gttcaagaaa	gaatcgtaaa	gcaaaagacg	tggtgcgga
2041	taagaagaagg	aagaaatagg	aagcgtcaaa	gcagatccac	gcattagaaa	atatcgaaag
2101	gggcccgggt	ctaaaagttc	ttacgttaga	acagtcaacc	gaggcgaatg	aaatgaattc
2161	gcagaagaaga	tatgggcagt	ctcctgtatt	tggttgcatt	gcgcgtcttg	agaaatgggtg
2221	gagtgctaga	aacgcaagcc	cggcttgtct	gcttaaaaga	gcctaccagg	ctaatgatcg
2281	cgatataatg	gataaaactg	aatggggaaa	agagatggg	tgagatctat	gttctgttac
2341	cgaagatgat	ctaaagggtt	ctaaagatga	ttctcatggt	tgagacatgt	ttattgtac
2401	accaaaagta	gcggctttca	aaggatcagc	ttcaatcaat	cttctggagt	gtctccatca
2461	agtctctcga	tgggcgcttg	ggctcggttg	gattttcttg	agtaggaatt	gtctattttg
2521	gtatgtgtat	ggaggtgggt	tgaatggctt	tgagcggttg	ttctacatta	actctgttgt
2581	ttaccctgtg	acagctctct	cgctcatcgt	ttactgttct	ctccctgcc	ctgtctctt
2641	cactggaaac	ttcatcgttc	ccgagattag	caactatcgc	agtatcctct	tcatggcgct
2701	ctctctcgta	attgcaataa	cgggttattc	cgagatgcaa	tgggggcaaa	tgatgtctga
2761	tgatgggttg	agaaacgaac	agttttgggt	catggagggt	gttctcgcgc	actgttttgc
2821	ttctctccaa	gttctctctc	aggttcttgc	tggtgtcgac	cttaacttca	cagtcacatc
2881	aaaaagcagc	gatgatggag	agttctctga	cccttaacct	ttcaaatgga	cttcaactct
2941	catccctcca	atgactctac	tcatacataa	gctcatgga	gtcatatgga	gagttcttga
3001	tgccatcagc	aatgggatac	actcgtgggg	accgcttttc	ggaaggactg	cttttgcact
3061	ttgggtcatt	attcatcctt	accggttctc	taaaaggttg	cttggaagaa	aagatgaagt
3121	gccaaacctt	attgtctgtc	gggtccatct	ctggcgctcg	attctctcac	ttctttgggt
3181	ccgggttaat	ccgttttgtg	cgaaaggcgg	ttcatattct	gagatctgtg	gtgtagacgt
3241	ctgtgatctc	gattgacccg	tggatgggtt	gggtgaaaaa	gtttcaattcc	cacggatcaa
3301	agagaggtaa	gagagataat	gttttacctc	taaaagactc	ctccatctgt	ttcattagat
3361	gatgaagatt	gaaagaaaaa	agaagattta	attttgttac	gagaatttgtt	atttttgcaa
3421	gaatgtgttg	tagatagcgg	ccgc			

10/10

M4481QF602DXS0131 WPD

FIG. 8

Arabidopsis thaliana cellulose synthase SEQ ID NO: 5

RPRLIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIE
LTVDGEPFVACNECAFPVCRPCYERREGNQACPOCKTRFKRLKGS PRVEGDEEEDD
IDDLNDEFYEGNNGIGFDQVSEGMSISRRNSGFFQSDLD SAPPGSQIPLITYGDEDVE
ISSDRHALIVPPSLGGHGNRVHPVSLSDPTVAHRRLMVPQKDLAVYGYGSVAWKDRM
EEWKRKQNEKLQVVRHEGDPDFEDGDDADF PMMDEGRQPLSMKIPKSSKINPYRMLI
VLRVLVILGLFFHYRILHPVKDAYALWLISV ICEIWFVAVSWLDDQFPKWYPIERETYL
RLSLRYEKEGKPSGLSPVDV FVSTVDPLKEPPLITANTVLSILAVDYPVDKRVACYVSD
DGAAMLTFEALSETAEFARKWVPFCKKYCIEPRAPEWYFCHKMDYLNKNVHPAFVRER
RAMKRDIYEEFKVKINALVATAQKVPEDGWTMQDGTWPFGNSVRDHPGMIQVFLGSDGV
RDVENNELPRLVVYSREKRPGFDHKKAGAMNSLIRVSGVLSNAPYLLNVDCDHYINN
SKALREAMCFMMDPQSGKKICYVQFPQRFDGIDRHDRYSNRNVVFFDINMKGDLGLQG
PIYVGTGCVFRRQALYGFDA P K K K K G P R K T C N C W P K W C L L C F G S R K N R K A K T V A A D K K
K K N R E A S K Q I H A L E N I E E G R G H K V L N V E Q S T E A M Q M K L Q K Y G Q S P V F V A S A R L E N G G
M A R N A S P A C L L K E A T Q V I S R G Y E D K T E W G K E I G W I Y G S V T E D I L T G S K M H S H G W R H V Y
C T P K L A A P K G S A P I N L S D R L H Q V L R W A L G S V E I F L S R H C P I W Y G Y G G L K W L E R L S Y I
N S V V Y P W T S L P L I V Y C S L P A I C L L T G K F I V P E I S N Y A S I L F M A L F S S I A I T G I E M Q W
G K V G I D D W W R N E Q F W V I G G V A H L F A L F Q G L L K V L A G V D T N F T V T S K A A D D G E F S D L Y
L F K W T S L L I P P M T L L I N V I G I V G V S D A I S N G Y D S W G P L F G R L F F A L W V I I H L Y P F L
K G L L G K Q D R M P T I I V V S I L L A S I L T L L W R V N P F V A K G G P I L E I C G L D C L

Declaration and Power of Attorney For Patent Application

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled METHOD OF ENHANCING CELLULOSE AND MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS (Attorney Docket No. 066040-9675), the specification of which was filed on May 18, 2000 as PCT International Application Number PCT/US00/13637, now filed under 35 U.S.C. 371 on November 19, 2001 as Application Serial No. 09/980,043.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims referred to above.

I acknowledge the duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

As a named inventor, I hereby appoint the following registered practitioners to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Customer Number

23409



TO: DIRECT ALL COMMUNICATIONS IN OR PERTAINING TO THIS APPLICATION

Joseph A. Gemignani
Michael Best & Friedrich LLP
100 East Wisconsin Avenue
Milwaukee, Wisconsin 53202-4108
(414) 271-6560

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of the foreign application for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

PCT/US00/13637	Prior Foreign Application	
(Number)	PCT	18 May 2000
	(Country)	(Day/Month/Year Filed)

I hereby claim priority benefit under Title 35, United States Code, §119 of the provisional U.S. patent application listed below:

<u>Application Serial No.</u>	<u>Filing Date</u>
60/135,280	21 May 1999

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first joint inventor: Vincent L. Chiang

Inventor's signature

Date:

Residence:

Citizenship:

Post Office Address:

Vincent L. Chiang
Nov. 19, 2001
Hancock, Michigan MI,
United States of America
1104 Birch Street
Hancock, MI 49930

Full name of second joint inventor: Luguang Wu

Inventor's signature

Date:

Residence:

Citizenship:

Post Office Address:

Luguang Wu
Dec 3, 2001
Kenmore, Australia
China
123 Akuna Street
Kenmore, QLD 4069, Australia

3-0 Full name of third joint inventor: Chandrashekhar P. Joshi

Inventor's signature Chandrashekhar P. Joshi

Date: 11/19/2007

Residence: Houghton, Michigan MI

Citizenship: India

Post Office Address: 713 Cedar Bluff Drive
Houghton, MI 49931

4-0 Full name of fourth joint inventor: Daniel T. Carraway

Inventor's signature Daniel T. Carraway

Date: 11-26-2007

Residence: Bainbridge, Georgia GA

Citizenship: United States of America

Post Office Address: 1910 Rich Street
Bainbridge, GA 31717

SEQUENCE LISTING

<110> Board of Control of Michigan Technological Univers

<120> METHOD FOR ENHANCING CELLULOSE AND MODIFYING LIGNIN BIOSYNTHESIS IN PLANTS

<130> 66040/9675

<140>

<141>

<150> 60/135,280

<151> 1999-05-21

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 3232

<212> DNA

<213> Populus tremuloides

<220>

<221> CDS

<222> (69)..(3002)

<400> 1

gtcgacccac gcgtccgtct tgaaagaata tgaagttgta aagagctggt aaagtggtaa 60

taagcaag atg atg gaa tct ggg gct cct ata tgc cat acc tgt ggt gaa 110

Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu

1 5 10

cag gtg ggg cat gat gca aat ggg gag cta ttt gtg gct tgc cat gag 158

Gln Val Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu

15 20 25 30

tgt agc tat ccc atg tgc aag tct tgt ttc gag ttt gaa atc aat gag 206

Cys Ser Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Asn Glu

35 40 45

ggc cgg aaa gtt tgc ttg cgg tgt ggc tgg cca tat gat gag aac ttg 254

Gly Arg Lys Val Cys Leu Arg Cys Gly Ser Pro Tyr Asp Glu Asn Leu

50 55 60

ctg gat gat gta gaa aag aag ggg tct ggc aat caa tcc aca atg gca 302

Leu Asp Asp Val Glu Lys Lys Gly Ser Gly Asn Gln Ser Thr Met Ala

65 70 75

tct cac ctc aac gat tct cag gat gtc gga atc cat gct aga cat atc 350

Ser His Leu Asn Asp Ser Gln Asp Val Gly Ile His Ala Arg His Ile

80 85 90

agt agt gtg tcc act gtg gat agt gaa atg aat gat gaa tat ggg aat 398

Ser Ser Val Ser Thr Val Asp Ser Glu Met Asn Asp Glu Tyr Gly Asn

WO 00/71670

PCT/US00/13637

95	100	105	110	
cca att tgg aag aat Pro Ile Trp Lys	cgg gtg aag agc Arg Val Lys Ser	tgt aag gat Cys Lys Asp	aaa gag aac aag Lys Glu Asn Lys	446
aag aaa aag aga agt Lys Lys Lys Arg	cct aag gct Pro Lys Ala	gaa act gaa cca gct Glu Thr Glu Pro	caa gtt cct Ala Gln Val Pro	494
aca gaa cag cag atg Thr Glu Gln Met	gaa gag aaa ccg Glu Glu Lys Pro	tct gca gag gct Ser Ala Glu Ser	tcg gag ccg Glu Glu Pro	542
ctt tca att gtt tat Leu Ser Ile Val	cca att cca cgc Pro Ile Pro Arg	aag aac ctc Lys Asn Lys	aca cca tac aga Leu Thr Pro Tyr	590
gca gtg atc att atg Ala Val Ile Ile	cga ctg gtc att Arg Leu Val Ile	ctg ggc ctc Leu Gly Leu Phe	ttc ttc cac ttc Phe Phe His Phe	638
aga ata aca aat cct Arg Ile Thr Asn	gtc gat agt gcc Pro Val Asp Ser	ttt ggc ctg tgg Phe Gly Leu Trp	ctt act tct Leu Thr Ser	686
gtc ata tgt gag atc Val Ile Cys Glu	tgg ttt gca ttt Trp Phe Ala Phe	tct tgg gtg ttg Ser Trp Val Leu	gat cag ttc Asp Gln Phe	734
ccc aag tgg aat cct Pro Lys Trp Asn	gtc aat aga gaa Pro Val Asn Arg	acg tat atc gaa Glu Thr Tyr Ile	agg ctg tcg Arg Leu Ser	782
gca agg tat gaa aga Ala Arg Tyr Glu	gag ggt gag cct Glu Gly Glu Pro	tct cag ctt gct Ser Gln Leu Ala	ggt gtg gat Gly Val Asp	830
ttt ttc gtg agt act Phe Phe Val Ser	gtt gat ccg ctg Val Asp Pro Leu	aag gaa ccg cca Lys Glu Pro Pro	ttg atc act Leu Ile Thr	878
gcc aat aca gtc ctt Ala Asn Thr Val	tcc atc ctt gct Leu Ser Ile Ala	gtg gac tat ccc Val Asp Tyr Pro	gtc gat aaa Glu Asp Lys	926
gtc tcc tgc tac gtg Val Ser Cys Tyr	tct gat gat ggt Val Ser Asp Asp	gca gct atg ctt Ala Ala Met Leu	tca ttt gaa Ser Phe Glu	974
tct ctt gta gaa aca Ser Leu Val Glu	gct gag ttt gca Ala Glu Phe Ala	agg aag tgg gtt Arg Lys Trp Val	ccg ttc tgc Pro Phe Cys	1022
aaa aaa ttc tca att Lys Lys Phe Ser	gaa cca aga gca Pro Arg Ala Pro	ccg gag ttt tac Glu Phe Tyr Phe	ttc tca cag Ser Ser Gln	1070

WO 00/71670

PCT/US00/13637

aaa att gat tac ttg	aaa gac aag gtt caa cct tct ttc gtg aaa gaa	1118
Lys Ile Asp Tyr Leu	Lys Asp Lys Val Gln Pro Ser Phe Val Lys Glu	
335	340 345 350	
cgt aga gca atg aaa agg gat tat gaa gag tac aaa gtc cga gtt aat	1166	
Arg Arg Ala Met Lys Arg Asp Tyr Glu Glu Tyr Lys Val Arg Val Asn		
355	360 365	
gcc ctg gta gca aag gct cag aaa aca cct gaa gaa gga tgg act atg	1214	
Ala Leu Val Ala Lys Ala Gln Lys Thr Pro Glu Glu Gly Trp Thr Met		
370	375 380	
caa gat gga aca cct tgg cct ggg aat aac aca cgt gat cac cct ggg	1262	
Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Thr Arg Asp His Pro Gly		
385	390 395	
cat gat tca ggt ctt cct tgg gaa ata ctg gga gct cgt gac att gaa	1310	
His Asp Ser Gly Leu Pro Trp Glu Ile Leu Gly Ala Arg Asp Ile Glu		
400	405 410	
gga aat gaa cta cct cgt cta gta tat gtc tcc agg gag aag aga cct	1358	
Gly Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro		
415	420 425 430	
ggc tac cag cac cac aaa aag gct ggt gca gaa aat gct ctg gtg aga	1406	
Gly Tyr Gln His His Lys Lys Ala Gly Ala Glu Asn Ala Leu Val Arg		
435	440 445	
gtg tct gca gta ctc aca aat gct ccc tac atc ctc aat gtt gat tgt	1454	
Val Ser Ala Val Leu Thr Asn Ala Pro Tyr Ile Leu Asn Val Asp Cys		
450	455 460	
gat cac tat gta aac aat agc aag gct gtt cga gag gca atg tgc atc	1502	
Asp His Tyr Val Asn Asn Ser Lys Ala Val Arg Glu Ala Met Cys Ile		
465	470 475	
ctg atg gac cca caa gta ggt cga gat gta tgc tat gtg cag ttc cct	1550	
Leu Met Asp Pro Gln Val Gly Arg Asp Val Cys Tyr Val Gln Phe Pro		
480	485 490	
cag agg ttt gat ggc ata gat aag agt gat cgc tac gcc aat cgt aac	1598	
Gln Arg Phe Asp Gly Ile Asp Lys Ser Asp Arg Tyr Ala Asn Arg Asn		
495	500 505 510	
gta gtt ttc ttt gat gtt aac atg aaa ggg ttg gat ggc att caa gga	1646	
Val Val Phe Phe Asp Val Asn Met Lys Gly Leu Asp Gly Ile Gln Gly		
515	520 525	
cca gta tac gta gga act ggt tgt gtt ttc aac agg caa gca ctt tac	1694	
Pro Val Tyr Val Gly Thr Gly Cys Val Phe Asn Arg Gln Ala Leu Tyr		
530	535 540	
ggc tac ggg cct cct tct atg ccc agc tta cgc aag aga aag gat tct	1742	
Gly Tyr Gly Pro Pro Ser Met Pro Ser Leu Arg Lys Arg Lys Asp Ser		
545	550 555	

WO 00/17670

PCT/US00/13637

tca tcc tgc ttc tca tgt tgc tgc ccc tca aag aag aag cct gct caa	1790
Ser Ser Cys Phe Ser Cys Cys Cys Pro Ser Lys Lys Lys Pro Ala Gln	
560	565
gat cca gct gag gta tac aga gat gca aaa aga gag gat ctc aat gct	1838
Asp Pro Ala Glu Val Tyr Arg Asp Ala Lys Arg Glu Asp Leu Asn Ala	
575	580
gcc ata ttt aat ctt aca gag att gat aat tat gac gag cat gaa agg	1886
Ala Ile Phe Asn Leu Thr Glu Ile Asp Asn Tyr Asp Glu His Glu Arg	
595	600
tca atg ctg atc tcc cag ttg agc ttt gag aaa act ttt ggc tta tct	1934
Ser Met Leu Ile Ser Gln Leu Ser Phe Glu Lys Thr Phe Gly Leu Ser	
610	615
tct gtc ttc att gag tct aca cta atg gag aat gga gga gta ccc gag	1982
Ser Val Phe Ile Glu Ser Thr Leu Met Glu Asn Gly Gly Val Pro Glu	
625	630
tct gcc aac tca cca cca ttc atc aag gaa gcg att caa gtc atc gcc	2030
Ser Ala Asn Ser Pro Pro Phe Ile Lys Glu Ala Ile Gln Val Ile Gly	
640	645
tgt ggc tat gaa gag aag act gaa tgg gga aaa cag att ggt tgg ata	2078
Cys Gly Tyr Glu Glu Lys Thr Glu Trp Gly Lys Gln Ile Gly Trp Ile	
655	660
tat ggg tca gtc act gag gat atc tta agt ggc ttc aag atg cac tgc	2126
Tyr Gly Ser Val Thr Glu Asp Ile Leu Ser Gly Phe Lys Met His Cys	
675	680
cga gga tgg aga tca att tac tgc atg ccc gta agg cct gca ttc aaa	2174
Arg Gly Trp Arg Ser Ile Tyr Cys Met Pro Val Arg Pro Ala Phe Lys	
690	695
gga tct gca ccc atc aac ctg tct gat aga ttg cac cag gtc ctc cga	2222
Gly Ser Ala Pro Ile Asn Leu Ser Asp Arg Leu His Gln Val Leu Arg	
705	710
tgg gct ctt ggt tct gtg gaa att ttc ttt agc aga cac tgt ccc ctc	2270
Trp Ala Leu Gly Ser Val Glu Ile Phe Phe Ser Arg His Cys Pro Leu	
720	725
tgg tac ggg ttt gga gga ggc cgt ctt aaa tgg ctc caa agg ctt gcg	2318
Trp Tyr Gly Phe Gly Gly Arg Leu Lys Trp Leu Gln Arg Leu Ala	
735	740
tat ata aac acc att gtg tac cca ttt aca tcc ctc cct ctc att gcc	2366
Tyr Ile Asn Thr Ile Val Tyr Pro Phe Thr Ser Leu Pro Leu Ile Ala	
755	760
tat tgc aca att cct gca gtt tgt ctg ctc acc gga aaa ttc atc ata	2414
Tyr Cys Thr Ile Pro Ala Val Cys Leu Leu Thr Gly Lys Phe Ile Ile	
770	775
cca acg ctc tca aac ctg gca agc atg ctg ttt ctt ggc ctc ttt atc	2462

WO 00/71670

PCT/US00/13637

Pro Thr Leu Ser Asn Leu Ala Ser Met Leu Phe Leu Gly Leu Phe Ile	
785 790 795	
tcc atc att gta act gcg gtg ctt gag cta aga tgg agc ggt gtc agc	2510
Ser Ile Ile Val Thr Ala Val Leu Glu Leu Arg Trp Ser Gly Val Ser	
800 805 810	
att gaa gat tta tgg cgt aat gaa caa ttc tgg gtg atc gga ggt gtt	2558
Ile Glu Asp Leu Trp Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Val	
815 820 825 830	
tca gcc cat ctc ttt gcg gtc ttc cag gga ttc tta aaa atg ttg gct	2606
Ser Ala His Leu Phe Ala Val Phe Gln Gly Phe Leu Lys Met Leu Ala	
835 840 845	
ggc atc gat acg aac ttc act gtc aca gca aaa gca gcc gaa gat gca	2654
Gly Ile Asp Thr Asn Phe Thr Val Thr Ala Lys Ala Ala Glu Asp Ala	
850 855 860	
gaa ttt ggg gag cta tat atg gtc aag tgg aca aca ctt ttg att cct	2702
Glu Phe Gly Glu Leu Tyr Met Val Lys Trp Thr Thr Leu Ile Pro	
865 870 875	
cca acc aca ctt ctc att atc aat atg tcg ggt tgt gct gga ttc tct	2750
Pro Thr Thr Leu Leu Ile Ile Asn Met Ser Gly Cys Ala Gly Phe Ser	
880 885 890	
gat gca ctc aac aaa gga tat gaa gca tgg ggg cct ctc ttt ggc aag	2798
Asp Ala Leu Asn Lys Gly Tyr Glu Ala Trp Gly Pro Leu Phe Gly Lys	
895 900 905 910	
gtg ttc ttt gct ttc tgg gtg att ctt cat ctc tat cca ttc ctt aaa	2846
Val Phe Phe Ala Phe Trp Val Ile Leu His Leu Tyr Pro Phe Leu Lys	
915 920 925	
ggc cta atg ggt cgc caa aac cta aca cca acc att gtt gtt ctc tgg	2894
Gly Leu Met Gly Arg Gln Asn Leu Thr Pro Thr Ile Val Val Leu Trp	
930 935 940	
tca gtg ctg ttg gcc tct gtc ttc tct ctc gtt tgg gtc aag atc aat	2942
Ser Val Leu Leu Ala Ser Val Phe Ser Leu Val Trp Val Lys Ile Asn	
945 950 955	
cca ttc gtt aac aaa gtt gat aac acc ttg gtt gcg gag acc tgc att	2990
Pro Phe Val Asn Lys Val Asp Asn Thr Leu Val Ala Glu Thr Cys Ile	
960 965 970	
tcc att gat tgc tgagctacct ccaataagtc tctcccagta ttttgggggtt	3042
Ser Ile Asp Cys	
975	
acaaacacct tgggaattgg aatatgatcc tcgttgtagt ttccctcaag aaagcacata	3102
tcgctgtcag tattttaaag aactgcaaga tgattgttct ctatgaagtt ttgaacagtt	3162
tgaaatgata ttatgttaaa atacaggttt tgattgtgtt gaaaaaaaaa aagaaaaaaaaa	3222

WO 00/71670

PCT/US00/13637

aaaaaaaaa

<210> 2
<211> 978
<212> PRT
<213> Populus tremuloides

<400> 2
Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu .
1 5 10
Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu .
20 25 30
Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Asn Glu .
35 40 45
Lys Val Cys Leu Arg Cys Gly Ser Pro Tyr Asp Glu Asn Leu .
50 55 60
Asp Val Glu Lys Lys Gly Ser Gly Asn Gln Ser Thr Met Ala .
65 70 75
Leu Asn Asp Ser Gln Asp Val Gly Ile His Ala Arg His Ile .
85 90
Val Ser Thr Val Asp Ser Glu Met Asn Asp Glu Tyr Gly Asn .
100 105 110
Trp Lys Asn Arg Val Lys Ser Cys Lys Asp Lys Glu Asn Lys .
115 120 125
Lys Arg Ser Pro Lys Ala Glu Thr Glu Pro Ala Gln Val Pro .
130 135 140
Gln Gln Met Glu Glu Lys Pro Ser Ala Glu Ala Ser Glu Pro .
145 150 155
Ile Val Tyr Pro Ile Pro Arg Asn Lys Leu Thr Pro Tyr Arg .
165 170
Ile Ile Met Arg Leu Val Ile Leu Gly Leu Phe Phe His Phe .
180 185 190
Thr Asn Pro Val Asp Ser Ala Phe Gly Leu Trp Leu Thr Ser .
195 200 205
Cys Glu Ile Trp Phe Ala Phe Ser Trp Val Leu Asp Gln Phe .
210 215 220
Trp Asn Pro Val Asn Arg Glu Thr Tyr Ile Glu Arg Leu Ser .
225 230 235
Tyr Glu Arg Glu Gly Glu Pro Ser Gln Leu Ala Gly Val Asp .
245 250

WO 00/71670

PCT/US00/13637

Val	Ser	Thr	Val	Asp	Pro	Leu	Lys	Glu	Pro	Pro	Leu	Ile	Thr	Ala	Asn
			260					265							270
Thr	Val	Leu	Ser	Ile	Leu	Ala	Val	Asp	Tyr	Pro	Val	Asp	Lys	Val	Ser
		275						280				285			
Cys	Tyr	Val	Ser	Asp	Asp	Gly	Ala	Ala	Met	Leu	Ser	Phe	Glu	Ser	Leu
		290				295					300				
Val	Glu	Thr	Ala	Glu	Phe	Ala	Arg	Lys	Trp	Val	Pro	Phe	Cys	Lys	Lys
305					310					315					320
Phe	Ser	Ile	Glu	Pro	Arg	Ala	Pro	Glu	Phe	Tyr	Phe	Ser	Gln	Lys	Ile
				325					330					335	
Asp	Tyr	Leu	Lys	Asp	Lys	Val	Gln	Pro	Ser	Phe	Val	Lys	Glu	Arg	Arg
			340					345					350		
Ala	Met	Lys	Arg	Asp	Tyr	Glu	Glu	Tyr	Lys	Val	Arg	Val	Asn	Ala	Leu
		355					360						365		
Val	Ala	Lys	Ala	Gln	Lys	Thr	Pro	Glu	Glu	Gly	Trp	Thr	Met	Gln	Asp
		370				375					380				
Gly	Thr	Pro	Trp	Pro	Gly	Asn	Asn	Thr	Arg	Asp	His	Pro	Gly	His	Asp
385					390					395					400
Ser	Gly	Leu	Pro	Trp	Glu	Ile	Leu	Gly	Ala	Arg	Asp	Ile	Glu	Gly	Asn
				405					410					415	
Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	Gly	Tyr
			420					425					430		
Gln	His	His	Lys	Lys	Ala	Gly	Ala	Glu	Asn	Ala	Leu	Val	Arg	Val	Ser
		435					440						445		
Ala	Val	Leu	Thr	Asn	Ala	Pro	Tyr	Ile	Leu	Asn	Val	Asp	Cys	Asp	His
		450				455					460				
Tyr	Val	Asn	Asn	Ser	Lys	Ala	Val	Arg	Glu	Ala	Met	Cys	Ile	Leu	Met
465					470					475					480
Asp	Pro	Gln	Val	Gly	Arg	Asp	Val	Cys	Tyr	Val	Gln	Phe	Pro	Gln	Arg
				485					490					495	
Phe	Asp	Gly	Ile	Asp	Lys	Ser	Asp	Arg	Tyr	Ala	Asn	Arg	Asn	Val	Val
			500					505					510		
Phe	Phe	Asp	Val	Asn	Met	Lys	Gly	Leu	Asp	Gly	Ile	Gln	Gly	Pro	Val
		515					520					525			
Tyr	Val	Gly	Thr	Gly	Cys	Val	Phe	Asn	Arg	Gln	Ala	Leu	Tyr	Gly	Tyr
		530				535					540				
Gly	Pro	Pro	Ser	Met	Pro	Ser	Leu	Arg	Lys	Arg	Lys	Asp	Ser	Ser	Ser
545					550					555					560

WO 00/71670

PCT/US00/13637

Gly Glu Leu Tyr Met Val Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr
865 870 875 880

Thr Leu Leu Ile Ile Asn Met Ser Gly Cys Ala Gly Phe Ser Asp Ala
885 890 895

Leu Asn Lys Gly Tyr Glu Ala Trp Gly Pro Leu Phe Gly Lys Val Phe
900 905

Phe Ala Phe Trp Val Ile Leu His Leu Tyr Pro Phe Leu Lys Gly Leu
915 920 925

Met Gly Arg Gln Asn Leu Thr Pro Thr Ile Val Val Leu Trp Ser Val
930 935 940

Leu Leu Ala Ser Val Phe Ser Leu Val Trp Val Lys Ile Asn Pro Phe
945 950 955 960

Val Asn Lys Val Asp Asn Thr Leu Val Ala Glu Thr Cys Ile Ser Ile
965 970 975

Asp Cys

<210> 3
<211> 1010
<212> DNA
<213> Populus tremuloides

<220>
<221> CDS
<222> (841)..(1008)

<220>
<223> 5' flanking region of PtCelA coding sequence

<400> 3
gaattcgccc ttttgaattc aggagacgat agtttccggt tcgttgaatg gctttgttca 60
cttctggtct agcaatttgc aaaagaagtt acaaaacaaa tgcataattat gtaaatttaa 120
caagagatgg gttctatggt cacttattta tgcccatcat ttgttctggg gttactcttt 180
atagtctgat tcgaagtgc aaactgccgt ttctggtatt gcaattatgt agccataaac 240
tgtaaatcct gtatctatta gcggaccaac aaccagatat acgggatcag cgtcgtaaaa 300
gagatctcca ttctacgttt ctttctaatt tttccgttcc agtgagagaa ttaccctgat 360
acattgacat gatgattgat gattatggga accattccga tgtagacac gagaccatct 420
ggatcctgcc agttttcagt tcacatggca tctcagccca agatcatgtg tttatacgcc 480
taatgacttg tattgaaagt ttggtaagtt gaagatgtgc tctgccaac agaaaccttc 540

WO 00/71670

PCT/US00/13637

```

cttaaaatttc cagcaaatct ttcaaacttg gccttacacc ccgaaaatag acgtgcttct 600
acttggggttc ttggaaacca tgcaccaacc gccatacccc accaaccacc caccctcaac 660
cttctctctcg ccattacaaa aatgtcagta ccaccctctg aaagacacca acacacccta 720
gctttgggta gggtatttga tataaaaaaaggccaaaac aaaagattgg aaggaagcag 780
aggaagaccc tcttgaaaga attgaagttg taaagagctg gtaaagtggg aataagcaag 840

atg atg gaa tct ggg gct cct ata tgc cat acc tgt ggt gaa cag gtg 888
Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu Gln Val
1 5 10 15

ggg cat gat gca aat ggg gag cta ttt gtg gct tgc cat gag tgt agc 936
Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu Cys Ser
20 25 30

tat ccc atg tgc aag tct tgt ttc gag ttt gaa atc aaa gag gcc cgg 984
Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Lys Glu Gly Arg
35 40 45

aaa gtt tgc ttg cgg tgt gcc tcg ag 1010
Lys Val Cys Leu Arg Cys Gly Ser
50 55

```

<210> 4
 <211> 56
 <212> PRT
 <213> Populus tremuloides
 <223> 5' flanking region of PtCelA coding sequence

```

<400> 4
Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu Gln Val
1 5 10 15

Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu Cys Ser
20 25 30

Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Lys Glu Gly Arg
35 40 45

Lys Val Cys Leu Arg Cys Gly Ser
50 55

```

<210> 5
 <211> 3444
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <223> cellulose synthase mRNA

<400> 5

WO 00/1670

PCT/US00/13637

gcggccgcg	ttaatcgccg	gttctcaca	caggaatgag	ttgtctctca	ttaatgccga	60
tgagaatgcc	cgaataagat	cagtcocaaga	gctgagtgga	cagacatgtc	aaatctgcag	120
agatgagatc	gaattgactg	ttgatggaga	accgtttgtg	gcatgtgaacg	aatgtgcatt	180
ccctgtgtgt	agacctgtgt	atgagtacga	aagacgagaa	ggcaatcaag	cttgtccaca	240
gtgcacaaacc	cggtttcaac	gtcttaaaagg	aagtccaaga	gttgaagggt	atgaagaagg	300
agatgacatt	gatgatttag	acaatgagtt	tgagtattgga	aaataaggga	ttggatttga	360
tcaggtttct	gaaggtatgt	caatctctcg	tcgcaactgc	ggtttccacc	aatctgattt	420
ggattcaegt	ccacctggct	ctcagattcc	attgctgact	tacggcgctc	aggacgttga	480
gattttctct	gatagacatg	ctctatttgt	tctctcttca	cttgtgtgag	atggccaatg	540
agttcatcct	gtttctctct	ctgaccgcag	cgtgggtgca	catcgaaagg	tgatgggtacc	600
tcagaaagat	cttgcggttt	atggttatgg	aagtgtcgct	tggaaaagatc	ggatggaggga	660
atggaagaga	aagcagaatg	agaaacttca	ggttgttagg	catgaaggag	atcctgattt	720
tgaagatggt	gatgatgctg	attttccaat	gatggatgag	ggaaggcagc	atgcattatg	780
gaagatacca	atcaaatcga	gcaagataaa	tccttaccgg	atgttaattg	tgctacgctt	840
tgtgattctt	ggtctctctt	ttcactaccg	tattcttcac	cccgctcaaa	atgcataatg	900
tttgtggctt	atttctgtta	tatgtgagat	atggtttgct	gtttctatgg	tctttgatca	960
gttccctaaa	tgttacccta	tcgagcgaga	gaccgtactt	gaccgatctc	atcgatgata	1020
tgagaaagaa	gggaaaccgt	cgggactatc	ccctgtggat	gtattttgta	gtcacagtga	1080
tccattgaaa	gagcctccgc	tattactgct	aaatactgtc	ttgtctattc	ttgctgttga	1140
ttatcctgtg	gataagggtt	cttgttacgt	atctgatgat	gggtcgtcta	tgcttacttt	1200
cgaaggctct	tctgagacgg	ctgaattcgc	aaggaaatgg	gttctcttca	gcaagaaata	1260
ttgtatttag	ctcgtgtctc	ccgaatggta	tttctgccat	aaaaatggac	acttgaagaa	1320
taaagtctcat	ccccattttg	ttagggagcg	gcgagccatg	aagagagatt	atgaagaatt	1380
caaagtaagg	atcaatgctt	tagtagcaac	agcacagaaa	gtgcctgagg	atggttggac	1440
tatgcaagag	gttacacctt	ggcccggtaa	tagtgtgcga	gatcatctcg	gcactgattca	1500
ggtcttctct	ggagtgtgag	gtgttctgtg	tgtcgaaaaac	aacgagtgct	ctcgattagt	1560
ttacgtttct	cggtgagaaga	gaccocggatt	tgatccacct	aagaaggctg	gagctatgaa	1620
ttccctgata	cggatctctg	gggtttctatc	aaatgctcct	tacctcttga	atgtcgattg	1680
tgactactac	atcaacaata	gcaaaagctc	tagagaagca	atgtgtttca	tgatggatcc	1740
tcagtacgga	aagaaaatct	gttatgttca	gttccctcaa	aggttcgatg	ggcttgatag	1800
gcacgatcga	tactcaaatc	gcaatgttgt	gttctttgat	atcaatgata	aaggttttga	1860
tgggtacaaa	ggccctatat	acgtcggtac	aggttgtgtt	ttcaggaagg	aagcgcctta	1920
cggatttgat	gcaccgaaga	agaagaaggg	cccacgtaag	acatgcaatt	gctggccaaa	1980
atggtgtctc	ctatgttttg	gttcaagaaa	gaatcgtaaa	gcaaagacag	tggtcgggga	2040
taagaagaag	aagaataggg	aagcgtcaaa	gcagatccac	gcattagaaa	atatcgaaag	2100
ggcccgcggt	ctaaaagttc	ttaacgtaga	acagtcaacc	gaggccaatg	aaatgaagtt	2160
gcagaagaaa	tatgggcagt	ctcctgtatt	tggtgcattc	gcgcgtctgg	agaatgggtg	2220
gatggctaga	aacgcaagcc	cggttgtctc	gcttaagaaa	gccatccaag	tcatttagtgc	2280
cggaatgaaa	gataaaactg	aatggggaaa	agagatttgg	tgaactatgt	gttctgttac	2340
cgaagatatt	cttaccgggt	ctaagatgca	tttctatggt	tggagacatg	tttatgtgat	2400
accaaaagtt	cggtctttca	aaggatcagc	tccaatcaat	cttcggagtc	gtctccatca	2460
agttcttcga	tgggcgcttg	ggtcggttga	gattttcttg	agtaggcatt	gtctctattt	2520
gtatggttat	ggaggtgggt	tgaattggct	tgcagcgggt	tctcacaatta	actcgttggt	2580
ttacccgctg	acctctctac	cgctcatcgt	ttactgttct	ctccctgccca	totgtcttct	2640
cactggaaaa	ttcatcgttc	ccgaattagg	caactatgca	agtaacctct	tcattgcgct	2700
ctctctcgtg	attgcaataa	cgggtattct	cgagatgcaa	tggggcaaa	ttgggattcga	2760
tgattggtgg	agaaacgaga	agttttgggt	cttgagaggt	tttctgcgc	atctgtttgc	2820
tctcttccaa	ggtctctcca	aggttcttgc	tggtgtcgac	actaaactca	cagtcacatc	2880
aaaagcagct	gatgatggag	agttctctga	cctttacctc	tccaatgaga	cttcaacttc	2940
catccctcca	atgacttctc	tcactataaa	cgtcatttga	gtcatagtga	gagtcctctg	3000
tgcacatcag	aatggatacg	actcgtgggg	accgcttttc	ggaagactgt	tctttgcaat	3060
ttgggtctac	attctctctt	accggttctc	taaaagtttg	cttggggaac	aagatagaat	3120
gccaaacctt	atttgtctct	ggtccactct	cctggcctcg	attcttaccac	ttctttgggt	3180
ccgggtttaa	ccgttttgtg	cgaaaggcgg	tctatttctc	gagatctctg	ttttagactg	3240
cttgtgattc	gattgaccgg	tggatgggtt	ggtgaaaaag	gttctaattc	cacgagatcaa	3300
agagaggtaa	gagagatatt	gttttacctc	taaaagactc	cttctattgt	ttcattagat	3360
gatgaaaaat	gaaaagaaaa	agaagattta	attttgttac	gagaatttgt	atttttgcaa	3420

WO 00/11670

PCT/US00/13637

gaatgtgttg tagatagcgg ccgc

3444

<210> 6
<211> 1080
<212> PRT
<213> Arabidopsis thaliana

<220>
<223> cellulose synthase

<400> 6
Arg Pro Arg Leu Ile Ala Gly Ser His Asn Arg Asn Glu Phe Val Leu
1 5 10 15
Ile Asn Ala Asp Glu Asn Ala Arg Ile Arg Ser Val Gln Glu Leu Ser
20 25 30
Gly Gln Thr Cys Gln Ile Cys Arg Asp Glu Ile Glu Leu Thr Val Asp
35 40 45
Gly Glu Pro Phe Val Ala Cys Asn Glu Cys Ala Phe Pro Val Cys Arg
50 55 60
Pro Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Asn Gln Ala Cys Pro Gln
65 70 75 80
Cys Lys Thr Arg Phe Lys Arg Leu Lys Gly Ser Pro Arg Val Glu Gly
85 90 95
Asp Glu Glu Glu Asp Asp Ile Asp Asp Leu Asp Asn Glu Phe Glu Tyr
100 105 110
Gly Asn Asn Gly Ile Gly Phe Asp Gln Val Ser Glu Gly Met Ser Ile
115 120 125
Ser Arg Arg Asn Ser Gly Phe Pro Gln Ser Asp Leu Asp Ser Ala Pro
130 135 140
Pro Gly Ser Gln Ile Pro Leu Leu Thr Tyr Gly Asp Glu Asp Val Glu
145 150 155 160
Ile Ser Ser Asp Arg His Ala Leu Ile Val Pro Pro Ser Leu Gly Gly
165 170 175
His Gly Asn Arg Val His Pro Val Ser Leu Ser Asp Pro Thr Val Ala
180 185 190
Ala His Arg Arg Leu Met Val Pro Gln Lys Asp Leu Ala Val Tyr Gly
195 200 205
Tyr Gly Ser Val Ala Trp Lys Asp Arg Met Glu Glu Trp Lys Arg Lys
210 215 220
Gln Asn Glu Lys Leu Gln Val Val Arg His Glu Gly Asp Pro Asp Phe
225 230 235 240

WO 00/71670

PCT/US00/13637

Glu	Asp	Gly	Asp	Asp	Ala	Asp	Phe	Pro	Met	Met	Asp	Glu	Gly	Arg	Gln
			245						250					255	
Pro	Leu	Ser	Met	Lys	Ile	Pro	Ile	Lys	Ser	Ser	Lys	Ile	Asn	Pro	Tyr
			260					265					270		
Arg	Met	Leu	Ile	Val	Leu	Arg	Leu	Val	Ile	Leu	Gly	Leu	Phe	Phe	His
		275					280				285				
Tyr	Arg	Ile	Leu	His	Pro	Val	Lys	Asp	Ala	Tyr	Ala	Leu	Trp	Leu	Ile
	290				295						300				
Ser	Val	Ile	Cys	Glu	Ile	Trp	Phe	Ala	Val	Ser	Trp	Val	Leu	Asp	Gln
305					310					315					320
Phe	Pro	Lys	Trp	Tyr	Pro	Ile	Glu	Arg	Glu	Thr	Tyr	Leu	Asp	Arg	Leu
				325					330				335		
Ser	Leu	Arg	Tyr	Glu	Lys	Glu	Gly	Lys	Pro	Ser	Gly	Leu	Ser	Pro	Val
			340					345					350		
Asp	Val	Phe	Val	Ser	Thr	Val	Asp	Pro	Leu	Lys	Glu	Pro	Pro	Leu	Ile
		355					360					365			
Thr	Ala	Asn	Thr	Val	Leu	Ser	Ile	Leu	Ala	Val	Asp	Tyr	Pro	Val	Asp
	370					375					380				
Lys	Val	Ala	Cys	Tyr	Val	Ser	Asp	Asp	Gly	Ala	Ala	Met	Leu	Thr	Phe
385					390					395					400
Glu	Ala	Leu	Ser	Glu	Thr	Ala	Glu	Phe	Ala	Arg	Lys	Trp	Val	Pro	Phe
				405					410					415	
Cys	Lys	Lys	Tyr	Cys	Ile	Glu	Pro	Arg	Ala	Pro	Glu	Trp	Tyr	Phe	Cys
			420					425					430		
His	Lys	Met	Asp	Tyr	Leu	Lys	Asn	Lys	Val	His	Pro	Ala	Phe	Val	Arg
		435					440					445			
Glu	Arg	Arg	Ala	Met	Lys	Arg	Asp	Tyr	Glu	Glu	Phe	Lys	Val	Lys	Ile
	450					455					460				
Asn	Ala	Leu	Val	Ala	Thr	Ala	Gln	Lys	Val	Pro	Glu	Asp	Gly	Trp	Thr
465					470					475					480
Met	Gln	Asp	Gly	Thr	Pro	Trp	Pro	Gly	Asn	Ser	Val	Arg	Asp	His	Pro
			485					490					495		
Gly	Met	Ile	Gln	Val	Phe	Leu	Gly	Ser	Asp	Gly	Val	Arg	Asp	Val	Glu
			500					505					510		
Asn	Asn	Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro
		515					520					525			
Gly	Phe	Asp	His	His	Lys	Lys	Ala	Gly	Ala	Met	Asn	Ser	Leu	Ile	Arg
	530					535					540				

WO 00/71670

PCT/US00/13637

Val	Ser	Gly	Val	Leu	Ser	Asn	Ala	Pro	Tyr	Leu	Leu	Asn	Val	Asp	Cys
545					550					555					560
Asp	His	Tyr	Ile	Asn	Asn	Ser	Lys	Ala	Leu	Arg	Glu	Ala	Met	Cys	Phe
			565						570					575	
Met	Met	Asp	Pro	Gln	Ser	Gly	Lys	Lys	Ile	Cys	Tyr	Val	Gln	Phe	Pro
			580					585					590		
Gln	Arg	Phe	Asp	Gly	Ile	Asp	Arg	His	Asp	Arg	Tyr	Ser	Asn	Arg	Asn
		595					600					605			
Val	Val	Phe	Phe	Asp	Ile	Asn	Met	Lys	Gly	Leu	Asp	Gly	Leu	Gln	Gly
	610					615					620				
Pro	Ile	Tyr	Val	Thr	Gly	Cys	Val	Phe	Arg	Arg	Gln	Ala	Leu	Tyr	Gly
625					630					635					640
Phe	Asp	Ala	Pro	Lys	Lys	Lys	Lys	Gly	Pro	Arg	Lys	Thr	Cys	Asn	Cys
				645					650					655	
Trp	Pro	Lys	Trp	Cys	Leu	Leu	Cys	Phe	Gly	Ser	Arg	Lys	Asn	Arg	Lys
			660					665					670		
Ala	Lys	Thr	Val	Ala	Ala	Asp	Lys	Lys	Lys	Asn	Arg	Glu	Ala	Ser	
		675					680					685			
Lys	Gln	Ile	His	Ala	Leu	Glu	Asn	Ile	Glu	Glu	Gly	Arg	Gly	His	Lys
	690					695					700				
Val	Leu	Asn	Val	Glu	Gln	Ser	Thr	Glu	Ala	Met	Gln	Met	Lys	Leu	Gln
705					710					715					720
Lys	Lys	Tyr	Gly	Gln	Ser	Pro	Val	Phe	Val	Ala	Ser	Ala	Arg	Leu	Glu
				725					730					735	
Asn	Gly	Gly	Met	Ala	Arg	Asn	Ala	Ser	Pro	Ala	Cys	Leu	Leu	Lys	Glu
			740					745					750		
Ala	Ile	Gln	Val	Ile	Ser	Arg	Gly	Tyr	Glu	Asp	Lys	Thr	Glu	Trp	Gly
		755					760					765			
Lys	Glu	Ile	Gly	Trp	Ile	Tyr	Gly	Ser	Val	Thr	Glu	Asp	Ile	Leu	Thr
	770					775					780				
Gly	Ser	Lys	Met	His	Ser	His	Gly	Trp	Arg	His	Val	Tyr	Cys	Thr	Pro
785					790					795					800
Lys	Leu	Ala	Ala	Phe	Lys	Gly	Ser	Ala	Pro	Ile	Asn	Leu	Ser	Asp	Arg
				805					810					815	
Leu	His	Gln	Val	Leu	Arg	Trp	Ala	Leu	Gly	Ser	Val	Glu	Ile	Phe	Leu
			820				825						830		
Ser	Arg	His	Cys	Pro	Ile	Trp	Tyr	Gly	Tyr	Gly	Gly	Gly	Leu	Lys	Trp
		835					840					845			

WO 00/71670

PCT/US00/13637

Leu Glu Arg Leu Ser Tyr Ile Asn Ser Val Val Tyr Pro Trp Thr Ser
850 855 860

Leu Pro Leu Ile Val Tyr Cys Ser Leu Pro Ala Ile Cys Leu Leu Thr
865 870 875 880

Gly Lys Phe Ile Val Pro Glu Ile Ser Asn Tyr Ala Ser Ile Leu Phe
885 890 895

Met Ala Leu Phe Ser Ser Ile Ala Ile Thr Gly Ile Leu Glu Met Gln
900 905 910

Trp Gly Lys Val Gly Ile Asp Asp Trp Trp Arg Asn Glu Gln Phe Trp
915 920 925

Val Ile Gly Gly Val Ser Ala His Leu Phe Ala Leu Phe Gln Gly Leu
930 935 940

Leu Lys Val Leu Ala Gly Val Asp Thr Asn Phe Thr Val Thr Ser Lys
945 950 955 960

Ala Ala Asp Asp Gly Glu Phe Ser Asp Leu Tyr Leu Phe Lys Trp Thr
965 970 975

Ser Leu Leu Ile Pro Pro Met Thr Leu Leu Ile Ile Asn Val Ile Gly
980 985 990

Val Ile Val Gly Val Ser Asp Ala Ile Ser Asn Gly Tyr Asp Ser Trp
995 1000 1005

Gly Pro Leu Phe Gly Arg Leu Phe Phe Ala Leu Trp Val Ile Ile His
1010 1015 1020

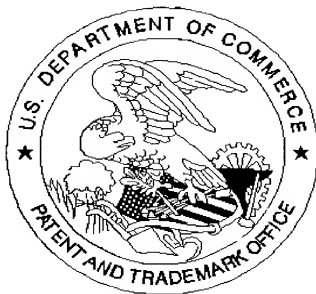
Leu Tyr Pro Phe Leu Lys Gly Leu Leu Gly Lys Gln Asp Arg Met Pro
1025 1030 1035 1040

Thr Ile Ile Val Val Trp Ser Ile Leu Leu Ala Ser Ile Leu Thr Leu
1045 1050 1055

Leu Trp Val Arg Val Asn Pro Phe Val Ala Lys Gly Gly Pro Ile Leu
1060 1065 1070

Glu Ile Cys Gly Leu Asp Cys Leu
1075 1080

United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ **Scanned copy is best available.**

Some drawing figures are too dark